California Renewables Portfolio Standard Renewable Generation Integration Cost Analysis

MULTI-YEAR ANALYSIS RESULTS AND RECOMMENDATIONS FINAL REPORT

CONSULTANT REPORT

Prepared For:
California Energy Commission

Prepared By:
California Wind Energy Collaborative

June 2006
CEC-500-2006-064

Prepared By:
California Wind Energy Collaborative
Henry Shiu
Davis, California
Contract No. 500-02-004
MR-017

PUBLIC INTEREST ENERGY RESEARCH
"Research Powers the Future"

Prepared For:
California Energy Commission

Michael Kane
Contract Manager

Dora Yen Nakafuji, Ph. D.
Project Manager

Elaine Sison-Lebrilla, P.E.
ENERGY GENERATION RESEARCH OFFICE

Martha Krebs, Ph. D.
Deputy Director
ENERGY RESEARCH AND DEVELOPMENT DIVISION

B.B. Blevins
Executive Director

California Renewables Portfolio Standard Renewable Generation Integration Cost Analysis

MULTI-YEAR ANALYSIS RESULTS AND RECOMMENDATIONS
FINAL REPORT

PREPARED FOR

The California Energy Commission

PREPARED BY

Henry Shiu
California Wind Energy Collaborative

Michael Milligan
National Renewable Energy Laboratory

Brendan Kirby
Oak Ridge National Laboratory

Kevin Jackson
Dynamic Design Engineering

DATE

June 2006

ACKNOWLEDGEMENTS

The authors are grateful to the California Energy Commission for their support of this study. We would also like to thank the California Independent System Operator, National Renewable Energy Laboratory, Oak Ridge National Laboratory, Southern California Edison, the Pacific Gas and Electric Company; and Analysis Group for their significant contributions of time and effort. In particular, we would like to recognize David Hawkins, Yuri Makarov, and Craig Taylor (CaISO); Gary Allen (SCE); Tom Miller and Linda Lea Weber (PG&E); Matthew Barmack and Ed Kahn (Analysis Group); Don Smith (CPUC/Office of Ratepayer Advocates); and Jim Lovekin (GeothemEx). Finally, we extend thanks to all those who have attended the workshops of this study or have otherwise provided valuable feedback, insight, and assistance.

Please cite this report as follows:

California Wind Energy Collaborative. 2006. *California Renewables Portfolio Standard Renewable Generation Integration Cost Analysis MULTI-YEAR ANALYSIS RESULTS AND RECOMMENDATIONS - FINAL REPORT.* California Energy Commission, PIER. Publication number CEC-500-2006-064.

PREFACE

The Public Interest Energy Research (PIER) Program supports public interest energy research and development that will help improve the quality of life in California by bringing environmentally safe, affordable, and reliable energy services and products to the marketplace.

The PIER Program, managed by the California Energy Commission (Energy Commission) conducts public interest research, development, and demonstration (RD&D) projects to benefit the electricity and natural gas ratepayers in California. The Energy Commission awards up to $62 million annually in electricity-related RD&D, and up to $15 million annually for natural gas RD&D.

The PIER program strives to conduct the most promising public interest energy research by partnering with RD&D organizations, including individuals, businesses, utilities, and public or private research institutions.

PIER funding efforts are focused on the following RD&D program areas:

- Buildings End-Use Energy Efficiency
- Industrial/Agricultural/Water End-Use Energy Efficiency
- Renewable Energy Technologies
- Environmentally Preferred Advanced Generation
- Energy-Related Environmental Research
- Energy Systems Integration

California Renewables Portfolio Standard Renewable Generation Integration Cost Analysis: MULTI-YEAR ANALYSIS RESULTS AND RECOMMENDATIONS is the final report for the California Renewables Portfolio Standard Renewable Generation Integration Cost Analysis, contract number 500-02-004, work authorization number MR-017, conducted by the California Wind Energy Collaborative (with assistance from the National Renewable Energy Laboratory, Oak Ridge National Laboratory, and Dynamic Engineering Design). The information from this project contributes to PIER's Renewable Energy Technologies program.

For more information on the PIER Program, please visit the Energy Commission's website at www.energy.ca.gov/pier or contact the Energy Commission at (916) 654-5164.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ACE	*Area Control Error*
ADS	*Automated Dispatch System*
AGC	*Automatic Generation Control*
CaISO	*California Independent System Operator*
CEMS	*Continuous Emissions Monitoring System*
CPUC	*California Public Utilities Commission*
CPS	*Control Performance Standard*
ELCC	*Effective Load Carrying Capability*
EPA	*Environmental Protection Agency*
EUE	*Expected Unserved Energy*
FERC	*Federal Energy Regulatory Commission*
Hz	*Hertz*
ICA	*Integration Cost Analyst*
IOU	*Investor Owned Utility*
ISO	*Independent System Operator*
kWh	*Kilowatt-hour (unit of energy)*
LCBF	*Least-Cost, Best-Fit*
LOLE	*Loss of Load Expectation*
LOLP	*Loss of Load Probability*
MW	*Megawatt (unit of power)*
MWh	*Megawatt-hour (unit of energy)*
MW-hr	*Megawatt of capacity committed for one hour*
NDA	*Nondisclosure Agreement*
NERC	*North American Electric Reliability Council*
NREL	*National Renewable Energy Laboratory*
ORNL	*Oak Ridge National Laboratory*
PG&E	*Pacific Gas and Electric Company*
PI	*Plant Information*
PIRP	*Participating Intermittent Resource Program*
PJM	*Pennsylvania-Jersey-Maryland Interconnection*

RPS	*Renewables Portfolio Standard*
RTO	*Regional Transmission Organization*
SCE	*Southern California Edison*

NOMENCLATURE

ACE	*area control error*
β	*control area frequency bias*
C_i	*capacity available in hour i*
ΔC_i	*effective capacity of analyzed resource at hour i*
ΔC_p	*effective capacity of analyzed resource at peak hour of year*
$COST_R$	*cost of regulation*
e_i	*scheduling error of generation resource i*
e_{Bias}	*system scheduling bias (difference between scheduled generation and forecasted load)*
$e_{Forecast}$	*load forecasting error*
$e_{Schedule}$	*system scheduling error (difference between scheduled generation and actual load)*
F_A	*actual system frequency*
F_S	*scheduled system frequency*
G	*total actual system generation*
g_i	*generation of analyzed resource i*
g_i	*generation of analyzed resource at hour i*
$g_{i,ave}$	*fifteen minute rolling average of generation of resource i*
$g_{i,lf}$	*load following component of generation resource i*
$g_{i,\mathrm{HA}}$	*hour-ahead scheduled generation of resource i*
$\overline{g_{i,15}}$	*fifteen minute average of generation of analyzed resource i*
I_{ME}	*meter error*
i	*generic index*
L	*total actual system load*
L_{actual}	*actual system load*
L_{ave}	*fifteen minute rolling average of system load*
$L_{HA\_Forecast}$	*hour ahead forecast of total system load*
$L_{HA\_Schedule}$	*hour ahead generation schedule*
L_i	*system load at hour i*
L_T	*total system compensation requirement*

$L_{T,ave}$	*fifteen minute rolling average of total system compensation requirement*
$\overline{L_{T,15}}$	*fifteen minute average of total system compensation requirement*
$\overline{L_{15}}$	*fifteen minute average of system load*
$LOLE$	*loss of load expectation*
$LOLE'$	*LOLE with resource of interest added to system*
N	*number of hours in the year*
NI_A	*actual net tie flows of control area*
NI_S	*scheduled net tie flows of control area*
P	*probability function*
R_{actual}	*actual amounts of purchased/self provided regulation*
R_i	*regulation requirement of analyzed resource*
$\hat{R}_i$	*allocated regulation share of analyzed resource*
$RATE_R$	*actual market rate of regulation*
r_i	*raw regulation component of analyzed resource*
r_L	*regulation component of total system load*
r_T	*regulation component of total system compensation requirement*
Δr_i	*regulation of system load less the resource of interest*
σ	*standard deviation*
σ_i	*standard deviation of regulation component of analyzed resource*
σ_T	*standard deviation of regulation component of total system compensation requirement*
σ_{T-i}	*standard deviation of regulation component of total system compensation requirement less the analyzed resource*
T	*total*
t	*time*
x	*dummy variable*

EXECUTIVE SUMMARY

Introduction

The California Renewables Portfolio Standard requires a “least-cost, best-fit” strategy for selecting new generation projects to fulfill its renewable energy supply goals. This explicitly includes indirect integration costs in the bid evaluation process. In previous work[2,3], integration costs were identified, valuation methodologies were defined, and a one year analysis of 2002 was performed.

Purpose

The purpose of this report is to document a multi-year analysis of integration costs, and apply the previously defined methodologies to a three year period from 2002 to 2004. The multi-year analysis provides opportunities to verify the consistency of the methodologies, further examine the practical issues associated with integration cost analysis, and to study the impact of renewables on integration costs over several years.

The methodologies are straightforward and were applied with little modification from their implementation in the previous one year analysis; the changes that were made are documented herein. The input data required for the analysis, however, was more problematic. Data quality and confidentiality issues hindered the progress of the study. The most critical data issues were ultimately resolved by using a combination of datasets from CaISO, SCE, and PG&E; performing extensive manual reviews of the data using custom developed programs; and training personnel who had access to the data to perform the analyses. However, outstanding data issues limited the analysis, as detailed within the report.

Results

Overall, the multi-year integration cost analysis results were reasonable, consistent with the analysis results of the previous one year dataset and, in some cases, verified with alternate approaches. The method used in this study for determining the capacity credit is the effective load carrying capability, ELCC. ELCC is a way to measure a power plant’s capacity contributions based on its impact to system reliability.

The results of the capacity credit analysis are summarized in the following table:

Resource	Capacity Credit					
	2002		2003		2004	
	ELCC relative to annual peak generation	ELCC relative to reported nameplate capacity	ELCC relative to annual peak generation	ELCC relative to reported nameplate capacity	ELCC relative to annual peak generation	ELCC relative to reported nameplate capacity
Medium Gas	100%	100%	100%	100%	100%	100%
Biomass	98%	98%	98%	98%	98%	98%
Geothermal (north)	108%	108%	109%	109%	109%	109%
Geothermal (south)	109%	109%	109%	109%	109%	109%
Solar	82%	88%	68%	83%	75%	79%
Wind (Northern Cal)	33%	24%	37%	25%	44%	30%
Wind (San Gorgonio)	42%	39%	28%	24%	27%	25%
Wind (Tehachapi)	29%	26%	34%	29%	29%	25%

The capacity credit analysis uses a conventional medium gas unit as a benchmark. Because of inconsistencies in the nameplate capacities provided for the generation aggregates, results are presented relative to both reported nameplate capacity and annual peak generation.

Biomass has outage rates comparable to the gas benchmark unit and, therefore, a high capacity credit. The geothermal outage rates are lower than the benchmark unit, resulting in a capacity credit exceeding 100%. The solar values are relatively high, as expected given its natural tendency to track load and the plants' auxiliary gas generators. Wind values ranged from 27% to 44% (based on annual peak generation; 24% to 39% based on reported nameplate capacity), with both regional and inter-annual variation. This is reasonable given wind's variable nature. The results were verified using an alternate method.

The results of the regulation analysis are summarized in the table below. Negative values indicate a cost.

Resource	Regulation Cost ($/MWh or mills/kWh)		
	2002	2003	2004
Total System	-0.42	-0.47	-0.39
Total Load	-0.41	-0.46	-0.36
Biomass	-0.09	-0.13	-0.12
Geothermal	-0.11	-0.03	-0.02
Solar	-0.44	-0.47	-0.37
Wind (Northern California)	-0.24	-0.40	-0.33
Wind (San Gorgonio)	-0.09	-0.43	-0.58
Wind (Tehachapi)	-0.57	-0.70	-0.56
Wind (Total)	-0.36	-0.53	-0.47

The resources studied have fairly minor impacts on total system regulation requirements. There is some inter-annual variation; in most cases, the changes follow the cost trend of actual regulation commitment by CaISO between 2002 and 2004. Because of the sheer size of total load, its regulation cost is consistently very close to that of the total system requirement. Geothermal, with a fairly flat output, has a low regulation cost, but a slightly higher value in 2002 when it was block scheduled for part of the year. The regulation costs of the solar and wind aggregates range between $0.24/MWh and $0.70/MWh, ignoring the anomalously low value for wind in San Gorgonio in 2002. While these values are higher than the results for biomass and geothermal, they are still quite modest. The solar results are consistent with the minute-to-minute variability in its generation data. The regulation costs imposed by wind are reasonable given that there are no apparent mechanisms that tie wind plant performance to the power system's needs either favorably or unfavorably in the regulation time frame.

The results of the load following analysis are summarized in the table below.

ERROR	2002				2003				2004			
	AVERAGE MINIMUM		AVERAGE MAXIMUM		AVERAGE MINIMUM		AVERAGE MAXIMUM		AVERAGE MINIMUM		AVERAGE MAXIMUM	
	MW	Compared to load forecast error alone (%)	MW	Compared to load forecast error alone (%)	MW	Compared to load forecast error alone (%)	MW	Compared to load forecast error alone (%)	MW	Compared to load forecast error alone (%)	MW	Compared to load forecast error alone (%)
Load forecast alone	-1945	100%	2112	100%	-1600	100%	2151	100%	-1439	100%	1529	100%
Load scheduling alone	-4747	244%	1302	62%	-4021	251%	2158	100%	-3700	257%	1776	116%
Scheduling bias	-5337	274%	1708	81%	-3336	208%	1534	71%	-3016	210%	1634	107%
Combined load forecast and renewable resource scheduling error												
Biomass	-1944	100%	2115	100%	-1603	100%	2157	100%	-1432	100%	1536	100%
Geothermal	-1947	100%	2112	100%	-1599	100%	2149	100%	-1442	100%	1529	100%
Solar	-1897	98%	2055	97%	-1631	102%	2153	100%	-1467	102%	1541	101%
Wind (Northern Cal)	-1946	100%	2148	102%	-1591	99%	2203	102%	-1419	99%	1554	102%
Wind (San Gorgonio)	-1930	99%	2142	101%	-1581	99%	2163	101%	-1443	100%	1545	101%
Wind (Tehachapi)	-1931	99%	2177	103%	-1569	98%	2181	101%	-1435	100%	1544	101%

The combined load forecast and renewable resource scheduling error values above indicate that renewables do not have a significant effect on the total energy requirements from the short term load following market at current penetration levels. The minimum scheduling bias was well over 200% greater than the combined forecast and scheduling error, implying that ample depth is available in the short term generator stack to handle incremental energy requirements.

A complementary methodology for analyzing ramping capability and requirements is also presented with a preliminary analysis. The ramping capability of thermal generators responding in the load following time frame appears to very large and capable of supporting a large amount of renewables. The ramping requirements of intermittent renewables appear to be significantly lower than the requirement of the total system load and the capability available in the CaISO control area.

At the conclusion of this analysis, a public workshop was held on April 3, 2006 to present and discuss the findings. The workshop generated additional comments from Pacific Gas and Electric Company and Southern California Edison. These comments, and a discussion of the issues they raise, are included as Appendices C and D. Some of the issues are beyond the scope of this analysis and some warrant further examination.

Recommendations

Provided the necessary data with sufficient quality, integration cost analysis becomes a relatively quick and straightforward process. An Integration Cost Analyst (ICA) is proposed to perform and report on integration cost analysis on a regular basis. It is recommended that the California Energy Commission or CPUC dedicate personnel and resources to perform the functions of an ICA. However, given the data issues encountered during this study, the tasks of handling/preparing data and analyzing integration costs should be made distinct and separate. This would also benefit other recent and current studies which require similar data. A data handling entity is proposed who would coordinate with data sources (CaISO, and IOUs) and the ICA to ensure the availability of good data quality as needed.

Benefits to California

The California Renewables Portfolio Standard challenges the state and its investor owned utilities to increase the amount of energy that is supplied from renewable sources. Meeting this challenge can reduce greenhouse gas emissions, moderate our dependence on natural gas, and mitigate the risks of electricity price volatility. Careful monitoring, reporting and analysis of the renewable energy data will help to provide Californians with the lowest cost and environmentally safest energy in the years ahead. This multi year integration cost analysis is a step forward in meeting this challenge.

1 INTRODUCTION

1.1 Policy Background

California has a large and diverse electric power supply network, which is critical for the economic and social well being of the state. In recent years, the California electric system was traumatized by a series of events that created power shortages, led to massive increases in the cost of electricity, caused the bankruptcy of Pacific Gas and Electric Company (PG&E), and led to severe financial hardship for the state's other Investor Owned Utilities (IOUs). One response to those dark times was the enactment of the Renewables Portfolio Standard (RPS, Senate Bill 1078)[1]. This law provides a means for improving supply diversity, while simultaneously reducing dependence on volatile fossil fuel resources. The primary goal of the RPS legislation is to expand and promote the economic use of California's abundant renewable energy resources.

California IOUs must supply an increasing portion of their energy mix from renewable energy sources, as a result of the RPS requirements. These energy sources are decoupled from traditional fuel markets and offer consistent pricing over long time periods, which are based primarily on capital recovery. California is blessed with significant renewable resources and remains a global leader in the application of these technologies. The state's renewable resource potential is more than sufficient to achieve the RPS goal of 20% renewable energy generation, although transmission capability constrains our ability to tap renewable energy in several key resource areas.

The RPS legislation envisioned annual procurements of new renewable resources through a bid selection process. Proposed renewable generation projects are expected to compete against one another to supply the IOUs with electricity, following a "least-cost, best-fit" (LCBF) process. The California Public Utilities Commission (CPUC) is charged with establishing and monitoring the LCBF process. According to the enabling legislation, the CPUC must:

> *"...adopt a process that provides criteria for the rank ordering and selection of least-cost and best-fit renewable resources to comply with the annual California Renewables Portfolio Standard Program obligations on a total cost basis. This process shall consider estimates of indirect costs associated with needed transmission investments and ongoing utility expenses resulting from integrating and operating eligible renewable energy resources."*

1.2 Overview of Study

This report documents a multi-year analysis of the integration costs of RPS eligible renewable resources. It is important to note that integration costs as discussed here are just a subset of potential indirect costs, which include investments in new transmission capacity and costs associated with remarketing electricity already purchased in long term supply contracts (Figure 1.1). As defined by statute, *integration costs* are the "*indirect costs associated with ongoing utility expenses from integrating* *and operating eligible renewable energy resources*." Other efforts have focused on transmission and remarketing costs; this report will discuss only methodologies and procedures recommended for calculating the indirect costs of integration.

The multi-year analysis is a derivative of the RPS Integration Cost Study, a California Energy Commission (Energy Commission) project to study integration costs in the context of RPS implementation. The overall goal of the project was to develop and define the procedures needed for the routine calculation of the indirect integration costs of RPS eligible renewable generators. The results obtained from those calculations are intended to support the CPUC's LCBF selection process of RPS bids.

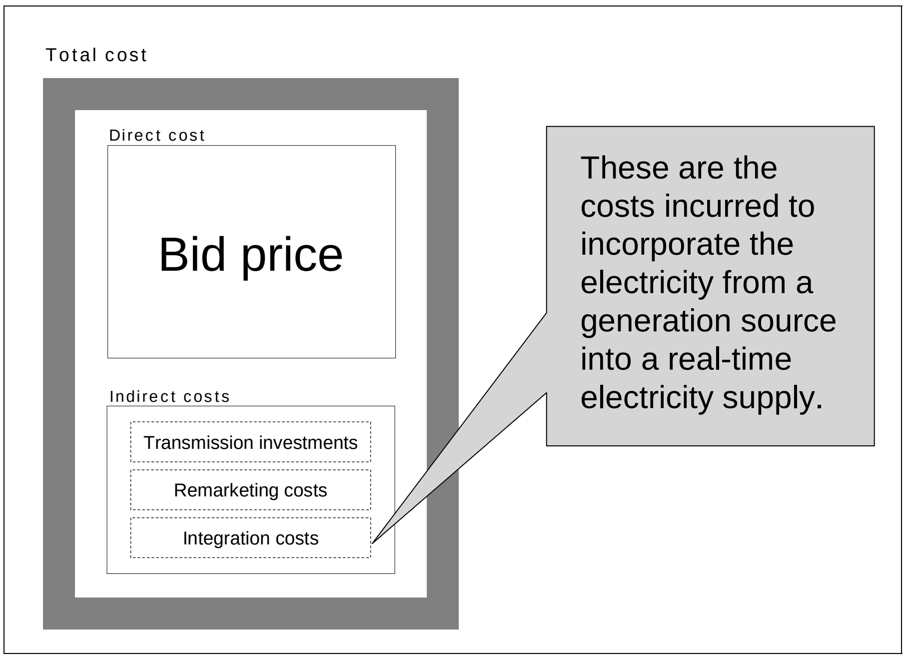

Figure 1.1 How integration costs fit in the least-cost, best-fit process.

The study was performed with a multi-phased approach. During Phase I, draft methodologies to quantify integration costs were presented in a public workshop conducted by the California Energy Commission in April of 2003. In mid-July, the California Independent System Operator (CaISO) provided a one year dataset containing electrical system and aggregated renewable generation data for 2002. Results of the analysis of this one year dataset and detailed methodology descriptions were presented in a draft report and an Energy Commission public workshop on 12 September 2003. Subsequently, public comments were reviewed and incorporated into a final draft of the Phase I report[2], which was published in December 2003. A second public workshop to review the Phase I findings and address public comments was held in February 2004. The final Phase I report was subsequently reviewed and adopted by the Energy Commission.

Phase II focused on studying the attributes of renewable generators that affect integration costs. This phase focused on geothermal and wind resources, which are expected to achieve the greatest penetration levels in the near term. Generator technologies and regional resource differences were identified and documented. Reports were completed and submitted to the Energy Commission in March of 2004.

Phase III presented a series of recommendations for the practical implementation of regular integration cost analysis. The final report[3], released in July 2004, proposed procedures for data handling and introduced an Integration Cost Analyst to regularly perform and report on integration cost calculations. Finalized methodologies were presented including revisions to the capacity credit analysis and results based on feedback from earlier workshops and the availability of improved hydro generation data. Phase III also discussed methods for studying the effect of generator attributes including different technologies and geographic regions. Finally, recommendations were made on how to apply the results of integration cost analyses to the RPS bid selection process. The Phase III findings were also presented in an Energy Commission public workshop in October of 2004.

The multi-year analysis documented in this report applies the integration cost valuation methodologies detailed in Phase III to a new multi-year dataset. The new analysis spans 2002 to 2004 and provided opportunities to verify the consistency of the methodologies and to further examine the practical issues associated with integration cost analysis. The methodologies were originally developed to be straightforward and were applied with little or no modification. They are detailed herein along with the analysis results. The methodologies, however, require good quality data and the difficulties encountered in assembling an adequate dataset hampered the analysis. Because these data issues will remain relevant to any future study, they are also detailed below. Finally, based on the experiences garnered from performing the multi-year analysis and resolving the data quality issues, recommendations are provided for future analyses.

2 CAPACITY CREDIT

2.1 Overview

Electricity is a unique commodity because it has two different units of value. Electric generation facilities provide energy value, but they also deliver capacity value. At any given time the power grid must have enough generating capacity to supply load demand. The system ultimately delivers energy to consumers, but without sufficient generating power the grid can become unstable and collapse into blackout. Power, or capacity, is critical to assure the reliability of the electric system. A generator's ability to deliver power when needed provides capacity value that is separate and distinct from the energy it delivers. The addition of new generating capacity will provide a value to the grid, because it increases system reliability during peak demand periods.

The value of capacity varies tremendously depending upon the system load and is highest when demand nears peak levels. For this reason it is important to understand the electrical demand patterns, which exhibit strong seasonal and diurnal trends. In this effort we reviewed data for statewide electrical power demand for a four-year period extending from 2001 to 2004. These data were sorted to determine the peak demand and the top twenty hours in each year are tabulated in Table 2.1. These data show that the months of July, September, and August are the most common peak demand periods, but that June can also have a very high load level.

Table 2.1. California peak demand hours for four years from 2001-2004. Times are in Pacific Standard Time.

2001		2002		2003		2004	
Date/Time	Demand (MW)	Date/Time	Demand (MW)	Date/Time	Demand (MW)	Date/Time	Demand (MW)
8/7 14:00	41155	7/10 13:00	42352	7/21 14:00	42581	9/8 14:00	45562
8/7 15:00	41017	7/10 12:00	41616	8/25 14:00	42506	9/8 15:00	45318
8/7 13:00	40493	7/9 14:00	41582	7/17 14:00	42502	9/7 14:00	45033
8/8 14:00	40488	7/9 13:00	41539	7/21 15:00	42346	9/8 13:00	44989
8/27 14:00	40439	7/9 15:00	41389	7/14 15:00	42227	9/9 14:00	44870
8/17 14:00	40384	7/10 14:00	41382	8/25 13:00	42218	9/7 15:00	44734
7/2 14:00	40241	9/23 14:00	41289	7/21 13:00	42184	8/11 14:00	44723
8/27 15:00	40173	6/5 14:00	41023	7/17 15:00	42143	9/9 15:00	44540
8/8 13:00	40149	6/5 15:00	40837	8/26 14:00	42107	8/11 15:00	44464
7/2 15:00	40073	9/23 15:00	40835	7/17 13:00	42037	8/10 14:00	44333
7/3 14:00	40065	7/10 15:00	40819	8/18 14:00	42007	8/10 15:00	44305
8/17 13:00	40017	9/3 13:00	40794	7/14 14:00	41968	7/21 14:00	44267
8/8 15:00	39953	8/9 14:00	40771	8/25 15:00	41905	8/11 13:00	44251
8/16 14:00	39900	8/12 14:00	40683	8/26 13:00	41826	9/10 14:00	44198
8/27 13:00	39899	7/12 14:00	40674	7/14 16:00	41655	7/20 14:00	44162
8/17 15:00	39847	7/9 12:00	40643	8/18 13:00	41613	7/21 15:00	44033
7/3 13:00	39741	7/12 13:00	40575	8/18 15:00	41433	9/7 13:00	44025
8/16 15:00	39733	9/23 13:00	40514	7/16 15:00	41412	7/20 15:00	43973
7/2 13:00	39690	6/5 13:00	40511	9/5 14:00	41394	9/9 13:00	43955
7/3 12:00	39650	8/12 15:00	40387	8/25 12:00	41368	7/19 14:00	43921

Although the selection of the top 20 hours is somewhat arbitrary, this table illustrates that the system peak in the ISO's annual system does not always occur in a given month. In fact, using this sample of 80 peak hours, we find that July and August dominate with 31 hours each, followed by September with 15 hours and June with 3 hours. To the extent that system risk (LOLP) is related to peak load, it appears that July and August are most important, but September can experience very high loads, based on this simple four-year sample. A later section of this chapter explores this relationship in more detail.

2.2 Definition of Capacity Credit

Renewable energy sources have operational characteristics that are different from conventional power generation facilities. One of the key differences is the intermittent production output of some renewable energy sources. Fortunately there are analytical methods for evaluating the capacity value of intermittent resources and correctly accounting for the value these generators provide to system reliability.

Evaluating the capacity provided by intermittent generators is more complicated than for conventional resources. The prior phases of this project used a reliability-based measure of capacity credit for all generators that were evaluated. The capacity credit of a specific generator is a function of the reliability of that generator, system demand, and other factors discussed below. No generator is perfectly reliable, so every type of generating resource has a capacity credit that is less than 100% of its maximum rated power. Some generators, because of decreased reliability or intermittent resource availability, will have a lower capacity credit than others.

Any generation resource that contributes to system reliability is providing capacity value and the preferred method for determining the capacity value is to calculate the *effective load carrying capability* (ELCC). This requires a reliability model that can calculate *loss of load probability* (LOLP), loss of load expectation (LOLE), or *expected unserved energy* (EUE). ELCC is a way to measure a power plant's capacity contributions based on its impact to system reliability. Using a measure such as ELCC, all power plants with a non-zero forced outage rate have an ELCC that is less than rated capacity (barring unusual plants with artificially low-rated capacity with respect to actual achieved capacity). The ELCC measure is often used as a way to compare alternative power plants, and can be easily applied to intermittent generators as well. A power plant's ELCC is typically calculated with an electric system reliability model or by a production-cost model.

The capacity credit represents the value of a generator's contribution to the reliability of the overall electrical supply system. In general the cost of capacity is determined using a benchmark technology, which is usually based upon natural gas. The relative capacity credit values were determined for various renewable technologies by comparing them to a combined cycle natural gas reference unit as the benchmark.

ELCC has been used for many years and can be applied to a wide variety of generators, not just renewables. This approach is well-grounded in electric power system reliability theory and applied methods. Although no generator has a perfect reliability index, we can use the concept as a benchmark to measure real generators. For example, a 500 MW generator that is perfectly reliable has an ELCC of 500 MW. If we introduce a 500 MW generator with a reliability factor of 0.85, or equivalently, a forced outage rate of 0.15, the ELCC of this generator *might* be 425 MW; however, the ELCC value cannot be calculated by simply multiplying the reliability factor by the rated plant output.

In general, the ELCC must be calculated by considering hourly loads and hourly generating capabilities. This procedure can be carried out with an appropriate production-simulation or reliability model. The electricity production simulation model calculates the expected loss of load. The usual formulation is based on the hourly estimates of LOLP, and the LOLE is the sum of these probabilities, converted to the appropriate time scale. The annual LOLE can be calculated as:

$$LOLE = \sum_{i=1}^{N} P\left[C_i < L_i\right] \qquad \textit{Equation 2.1}$$

where *P()* denotes the probability function, *N* is the number of hours in the year, C_i represents the available capacity in hour *i*, and L_i is the hourly utility load. To calculate the additional reliability that results from adding intermittent generators, we can write *LOLE'* for the *LOLE* after renewable capacity is added to the system as:

$$LOLE' = \sum_{i=1}^{N} P\left[(C_i + g_i) < L_i\right] \qquad \textit{Equation 2.2}$$

where g_i is the power output from the generator of interest during hour i. The ELCC of the generator is the additional system load that can be supplied at a specified level of risk (loss of load probability or loss of load expectation).

$$\sum_{i=1}^{N} P(C_i < L_i) = \sum_{i=1}^{N} P\left[(C_i + g_i) < (L_i + \Delta C_i)\right] \qquad \textit{Equation 2.3}$$

Calculating the ELCC of the renewable plant amounts to finding the values ΔC_i that satisfy Equation 2.3. This equation says that the increase in capacity that results from adding a new generator can support ΔC_i more MW of load at the same reliability level as the original load could be supplied (with C_i MW of capacity). To determine the annual ELCC, we simply find the value ΔC_p, where p is the hour of the year in which the system peak occurs after obtaining the values for ΔC_i that satisfy the equation. Because LOLE is an increasing function of load, given a constant capacity, we can see from Equation 2.3 that increasing values of ΔC_i are associated with declining values of *LOLE*. Unfortunately, it is not possible to analytically solve Equation 2.3 for ΔC_p. The solution for ΔC_p involves running the model for various test values of ΔC_p until the equality in Equation 2.3 is achieved to the desired accuracy.

Although the level of detail of the input data varies between models, hourly electric loads and generator data are required to calculate LOLE. Common outputs from these models include various costs and reliability measures, although cost data are not used to perform system reliability calculations. Some of the models used for these calculations are chronological, and others group related hours to calculate a probability distribution that describes the load level.

2.3 Methodology and Analysis Description

2.3.1 STEP-BY-STEP ELCC BASED CAPACITY CREDIT ANALYSIS METHODOLOGY

The ELCC modeling requires a production/market simulation or reliability model that is capable of representing the California power supply system and calculating LOLP, LOLE, or other similar reliability metric. For each intermittent resource, and for hydro and interchanges, hourly production should be used. The overall approach is to run the model with all generators included and adjusting loads so that a target reliability level is met. This is often one day in ten years LOLE, but could be another reliability target if desired. The renewable generator is then replaced with varying levels of a benchmark unit. In the Phase I work we used a combined cycle natural gas unit as the primary benchmark. The benchmark could also be a simple combustion turbine, if that unit used to determine the cost of capacity. When a given quantity of gas brings the annual LOLE back to the reliability target, the quantity of gas is noted, and is the ELCC of the renewable generator. The detailed step-by-step approach is as follows:

Table 2.2. Step-by-step description of ELCC based capacity credit analysis methodology.

1.	Develop a time series that represents hourly generation of the candidate resource.
2.	For existing intermittent renewables, develop a similar time series, one for each renewable.
3.	Add these resources to the supply model of the California system.
4.	Run the reliability model.
5.	Note the annual loss of load expectation. We want a target of 1 day/10 years, which equates to 2.4 hours/year LOLE. It is unlikely that we will obtain our target 1 day/10 years in this initial run. The reliability metric is sometimes (erroneously) displayed as annual LOLP by the model.
6.	Adjust the hourly loads, if necessary. If the LOLE exceeds 2.4 hours/year (this is highly unlikely in the base case) then pro-rate the hourly loads downward and rerun the model. If the LOLE is less than 2.4 hours/year, then pro-rate the hourly loads upward and rerun the model. Continue repeating steps 4-6 until the reliability target has been met.
7.	The final modeling run from step 6 is the base case, and represents the reliability target of 1 day/10 years LOLE. Save this load set.
8.	Remove the renewable generator of interest. Although not strictly necessary, you can rerun the model at this point. If the model is run, the reliability will decrease (LOLE will increase).
9.	Incrementally add the gas benchmark unit. If the reliability model makes it easy to run alternative, multiple scenarios, the gas benchmark unit can be added incrementally in a batch of modeling runs. Alternatively, some models allow the user to specify a target output and a “rule” for changing inputs so that the goal is reached. In any case, each incremental addition of the reference unit will result in a new annual LOLE value. At each of these steps, the model should save total gas capacity for this step and the annual LOLE. This set of runs must add sufficient gas capacity to bring the LOLE down to the benchmark reliability level of 1 day/10 years, or lower. The results of these iterative steps can be inserted in a spreadsheet.
10.	The ELCC of the generator of interest is the gas capacity that corresponds to the case that matches the original reliability target.

2.3.2 ANALYSIS CHANGES FROM PHASE I AND PHASE III

Based on discussion with utilities and other stakeholders throughout the RPS Integration Cost project, several refinements were made to the ELCC calculation of renewable technologies. The Phase I report modeled the renewable intermittent generation using a probabilistic approach. This method is similar to what is often done with conventional units that have multiple output settings, each with an associated partial forced outage rate. As a result of extensive feedback during public workshops, the probabilistic method was replaced with a more direct approach that uses actual hourly output of the renewable generators.

The probabilistic approach is more appropriate as an indicator of future performance, where there are considerable uncertainties surrounding the timing of the power delivery from certain resources. Directly using hourly output is more appropriate for measuring past contributions to capacity from an intermittent resource. It does not consider alternative timing of the power delivery from intermittent resources, as does the probabilistic method. However, when multiple years of data are analyzed, this is not a significant limitation. Therefore, single-year estimates should be considered as such, and would be expected to vary somewhat from year to year. This was discussed in detail and applied in the Phase III update to the one year capacity credit study. In the multi-year study, we continued to use the direct hourly method.

Other improvements were made in the input data. For the multi-year analysis we utilized renewable generation data directly from the IOUs. This allowed us to bypass some data from CaISO's Plant Information (PI) system that suffered from data errors. Those errors were sometimes difficult to detect because the renewable generation data was aggregated, which tended to obscure the errors. The data errors caused artificial offsets to actual generation and injected unrealistic ramping behavior over long time periods into the data set. The CaISO data also had related problems with the reported nameplate capacity of the generator aggregates. The IOU data aggregates used for the multi-year analysis were the ones that most closely matched the CaISO data used in the regulation and load following analyses, below. The input datasets are discussed in detail in Section 4.4.

We were able to obtain one-minute hydro data from CaISO and used hourly averages of this data directly in the multi-year analysis. This is an improvement over the hydro modeling previously used. In Phase III, an optimal dispatch of hydro was used based on California Energy Commission information on monthly minimum and maximum flows and rough estimates of pond-storage and pumped hydro data. However, a significant portion of hydro energy is run-of-river, which is uncontrollable and subject to nature. This is similar to wind and solar, although hydro is less variable than wind and has different characteristics than solar. But ultimately, these forms of generation are not dispatchable. As discussed further in some of the workshops and the Phase III report, the impact of the hydro system on the hourly risk profile is significant. The results below support this view and also show the significant effect of the interchange.

The outcome of the public workshops during the Phase I work suggested that scheduled maintenance from conventional units should be eliminated from the modeling and was excluded in the one year and multi-year analyses. As we stated in the Phase III report, whether this should continue is a policy question. Workshop participants in the earlier phases of this project suggested that in principle, the capacity value of renewable generators should be independent from conventional maintenance scheduling.

2.4 Multi-Year Analysis Results and Discussion

2.4.1 RELATIONSHIP BETWEEN RELIABILITY, LOAD, INTERCHANGE, AND HYDRO

Power systems experience a wide variety of conditions from year to year. Because load is generally sensitive to weather, unusually warm or cool temperatures can cause the load profile in a given year to diverge from “normal.” Generation does not always respond in the same way to nearly identical load conditions. Because loads can change significantly from year to year, both in magnitude and timing, one would expect that reliability indicators such as LOLP would also change, perhaps significantly. Because LOLP is a key ingredient in calculating capacity credit, we began the analysis by collecting the results of the base case reliability model runs for each of the three year periods (as discussed below, note that 2004 is represented by data from September 2003 to September 2004). Figure 2.1 is a LOLP-duration curve for each year, plotted on the same graph. We can see from the graph that 2004 exhibits a relatively sharp decline in LOLP as loads drop off from the annual peak. Much of the annual risk occurs in a smaller number of hours, whereas the curves for 2002 and 2003 indicate a more gradual decline. In 2002 the risk is spread over more hours. The significance of this graph is that the risk profile of the CaISO system, as measured by LOLP, changes from year to year. It is not possible *a priori* to determine which hours will have the highest risk, or even to predict the risk profile with certainty.

For a closer view, we generated a series of graphs for the three-year period that show not only the relationship between load and LOLP, but the overall impact that the hydro system and interchange have on risk. In general (ignoring hydro and interchange), the highest annual LOLP would be expected to occur during the peak hour. However, there are many factors that can cause LOLP in near-peak hours to exceed the LOLP on the system peak. Generator schedules, exchange schedules, and hydro generation are capable of responding to the high prices that accompany peak or near-peak loads, subject to operating constraints. It is therefore possible that real-time reserves are higher during system peak than at near-peak. These and other factors can contribute to a LOLP profile that is similar to, but does not match, the peak load profile.

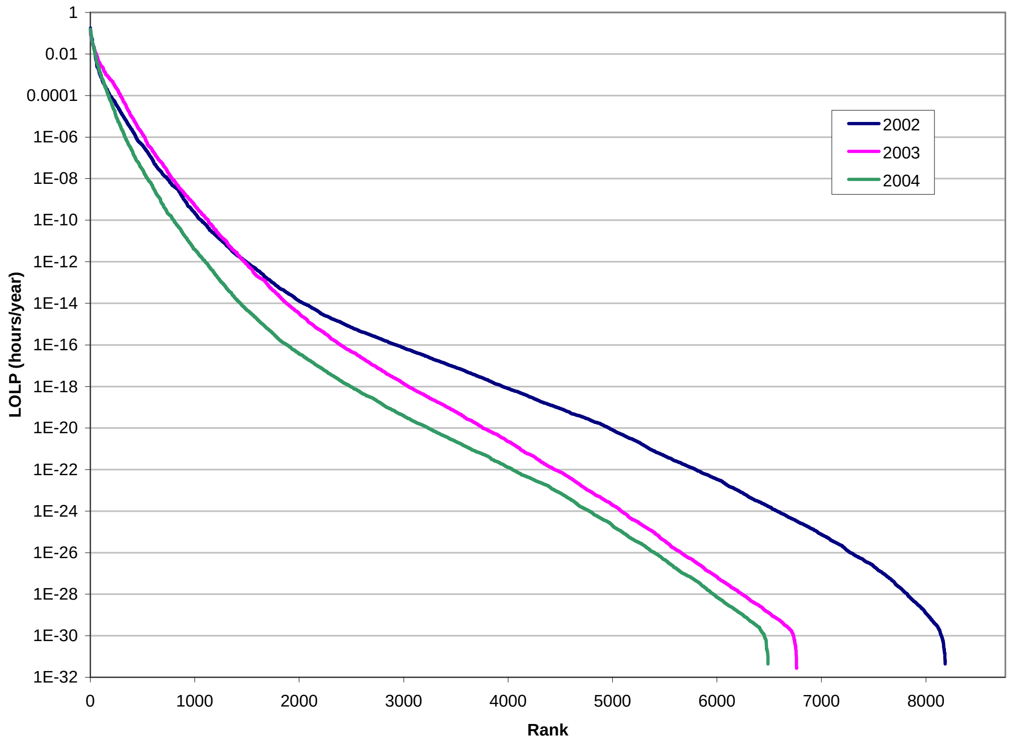

Figure 2.1. Hourly LOLP, ranked, 2002-2004.

In Figure 2.2, the graph shows a typical load duration curve, in this case for 2002 (the blue line with the smooth characteristic). Superimposed on this graph are two additional rankings. The first shows the ranking of load by hourly LOLP (red). What the graph shows is that high load hours may generally be correlated with high LOLP, but the correlation is weak when we view the top 271 hours (the somewhat arbitrary cutoff point was LOLP >= 0.000001 days/year).

The final ranking on the graph (green) is based on the load that remains to be served after hydro and interchange have been taken into account. We refer to this as the load, net of hydro and interchange. Because hydro's and imports' forced outage rates are very low and/or cannot be objectively assessed, standard practice is to ignore forced outage rates for these resources. The implication is that the primary impact that hydro and imports have on system risk is to shift the timing of risk. For intermittent resources such as wind and solar (ignoring gas-assist for the moment) this further implies that for the generator to reduce annual LOLE, it must provide power during periods of high LOLP after taking account of hydro and imports/exports. This can have a significant impact on the LOLP profile, which is apparent from the figure.

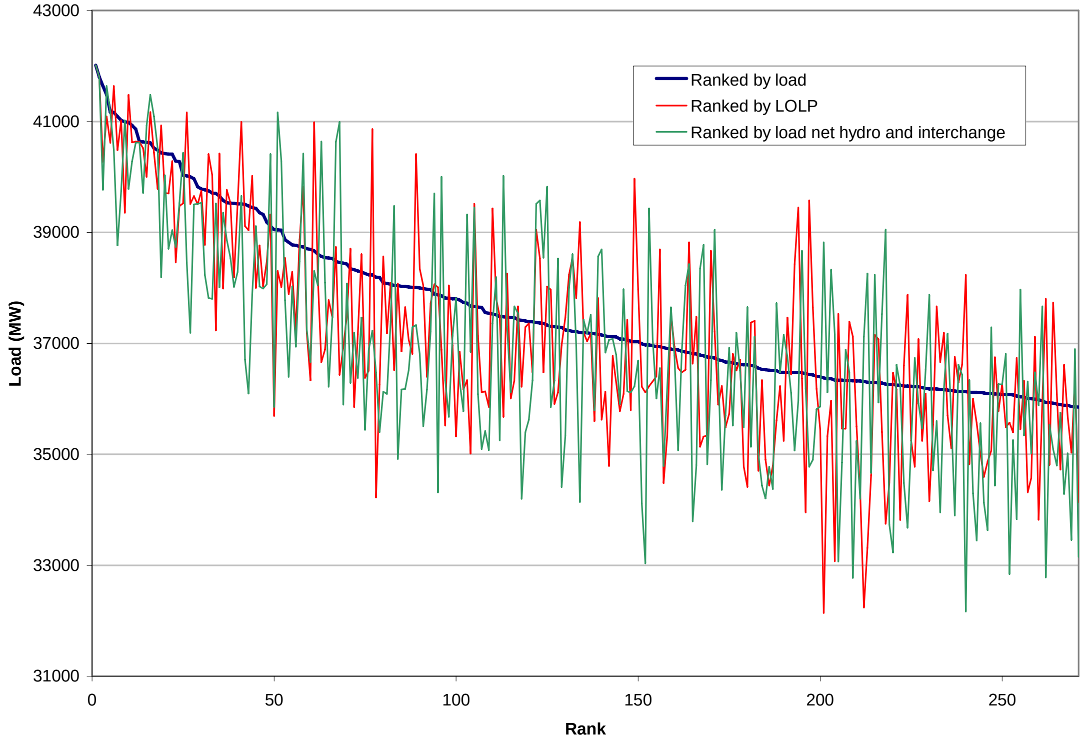

Figure 2.2. Load in 2002 during top risk hours, ranked by load, LOLP, and load net hydro and interchange.

Figure 2.3 takes a closer look at the load net hydro and interchange. The LOLP duration curve is not monotonically decreasing as a function of net load. If an intermittent resource delivers its energy during the high LOLP events, it will achieve a relatively high capacity credit. The timing of these high LOLP events will not necessarily correspond to highest load events.

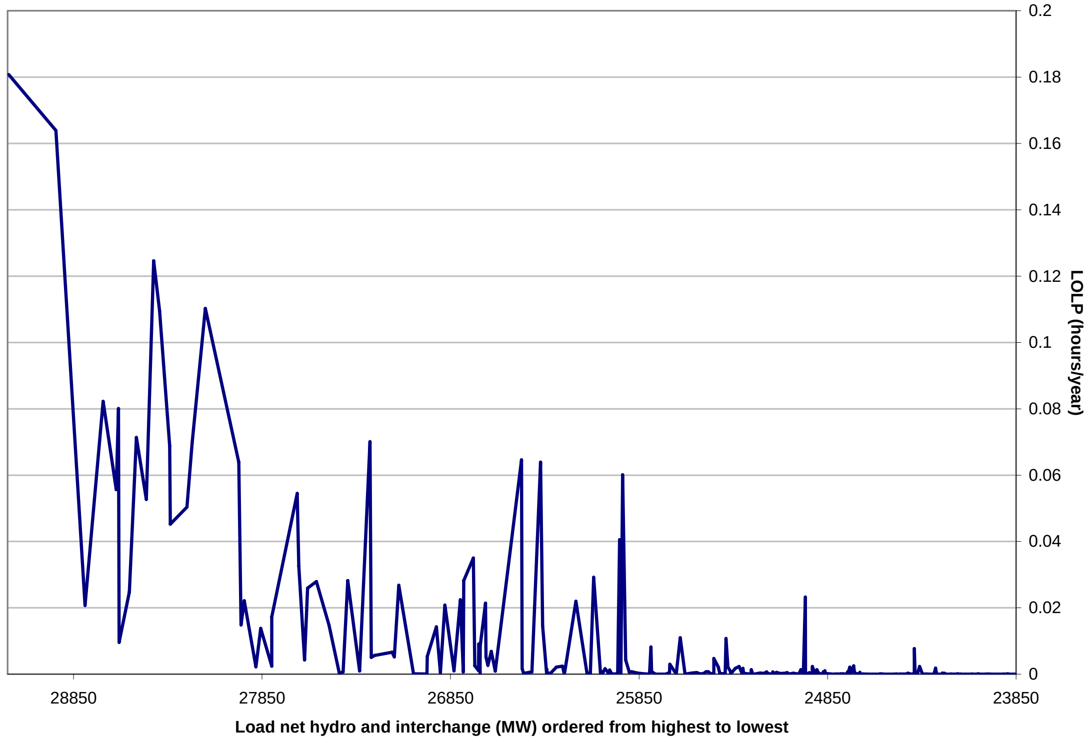

Figure 2.3. LOLP in 2002 at top hours of load net hydro and interchange.

The following series of four figures show similar characteristics in 2003 and 2004.

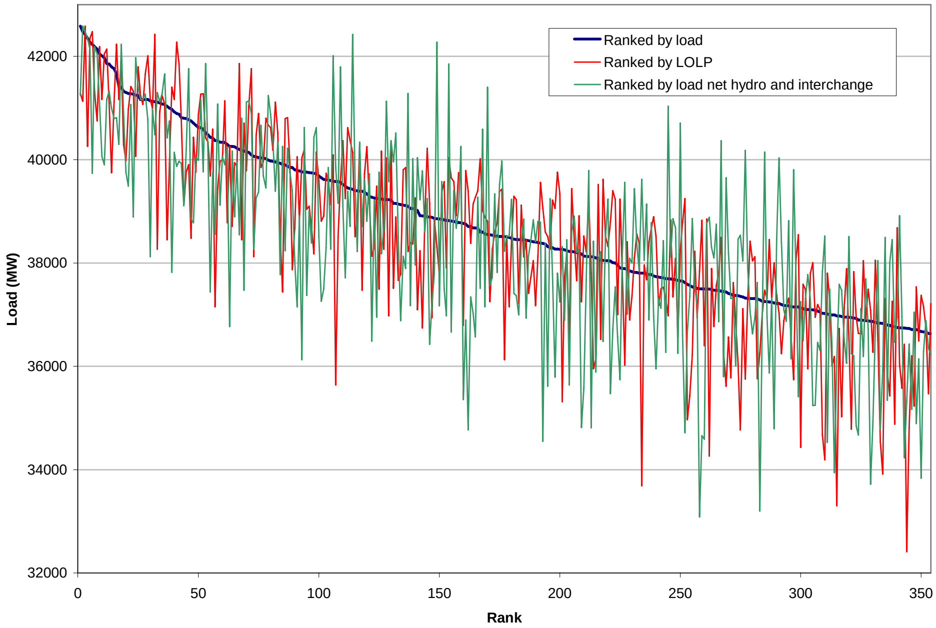

Figure 2.4. Load in 2003 during top risk hours, ranked by load, LOLP, and load net hydro and interchange.

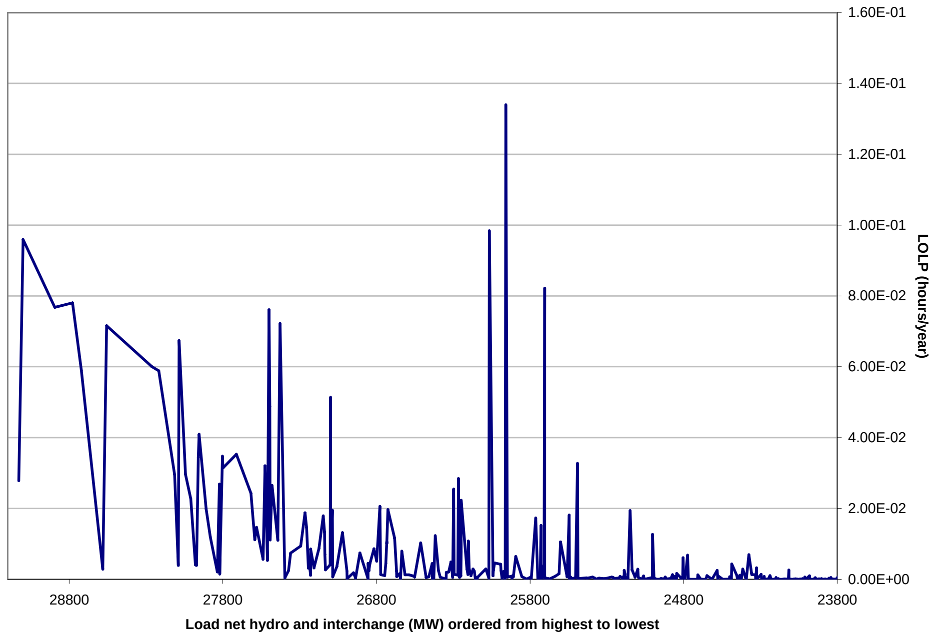

Figure 2.5. LOLP in 2003 at top hours of load net hydro and interchange.

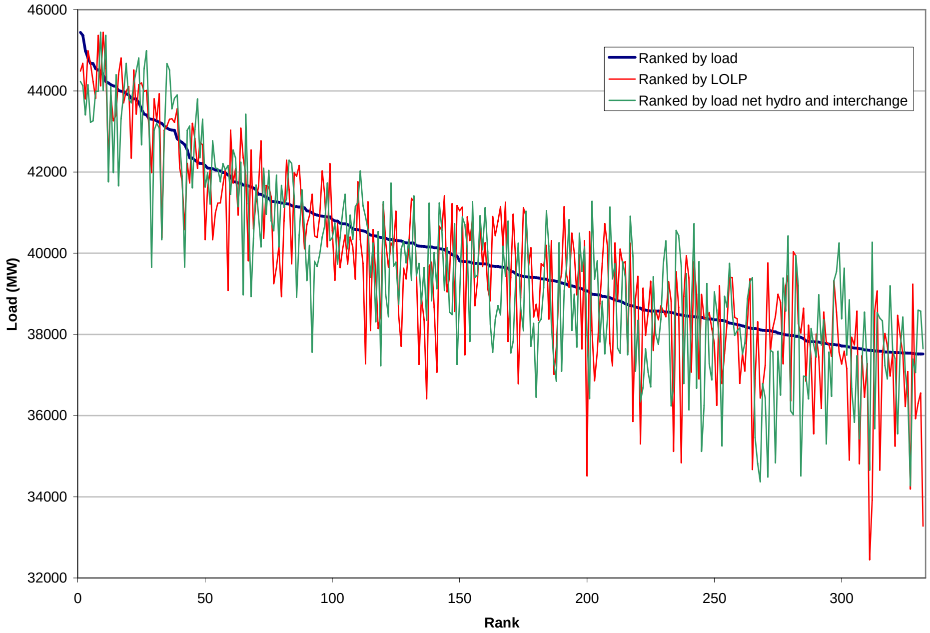

Figure 2.6. Load in 2004 during top risk hours, ranked by load, LOLP, and load net hydro and interchange.

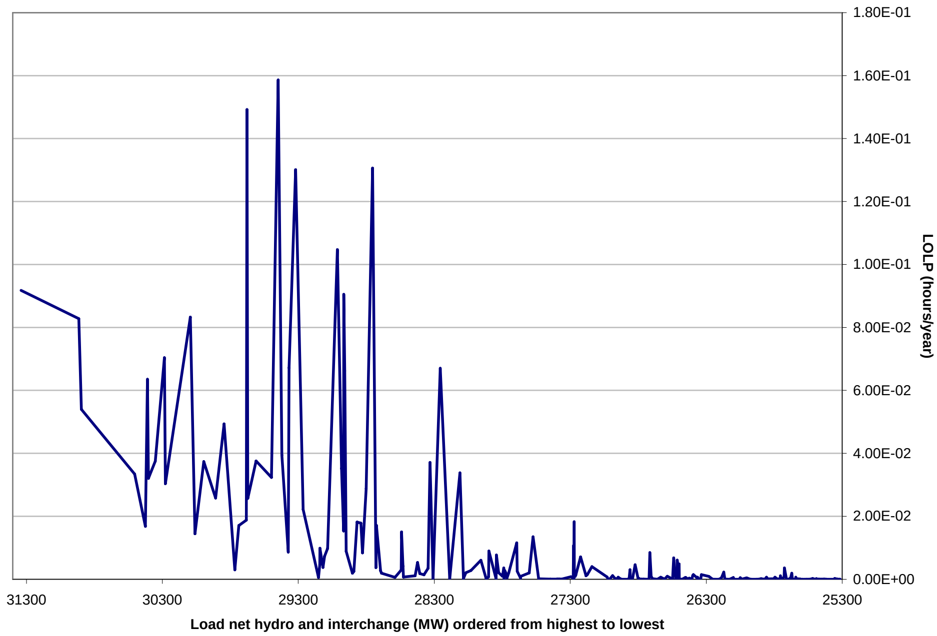

Figure 2.7. LOLP in 2004 at top hours of load net hydro and interchange.

2.4.2 ELCC RESULTS

All ELCC results were calculated based on the method outlined above, and with data received from CaISO and the IOUs. The analysis requires a complete year of data for each calculation. The input datasets used had complete years for 2002 and 2003, but not 2004. 2004 is represented by data from mid-September 2003 to mid-September 2004. It is simply referred to as “2004” for convenience.

As in the prior work, ELCC is measured relative to a benchmark unit, a gas combined cycle generator with a 4% forced outage rate and a 7.6% annual maintenance rate. The biomass and geothermal resources were modeled as conventional generators. Because of the very low forced outage rates for geothermal units (0.66%) and low maintenance rates (2.61%) geothermal plants are able to provide more capacity value than the benchmark units. The biomass generation was modeled with a 5.15% forced outage rate and 7.91% maintenance rate. All wind and solar resources were modeled as time series, using the actual hourly generation provided by the IOUs for the full year. Transactions (interchange) and hydro were also represented by actual hourly data, obtained from CaISO. We note that in February 2002 there were some errors in the hydro data, which we patched through a combination of interpolation and pattern matching. Because LOLP during the month of February is so close to zero, this will not impact the results.

During the processing of the data for the analysis, some discrepancies were uncovered in the reported nameplate capacities of some of the generation aggregates. In prior work we reported capacity value as a percent of the annual maximum hourly generation for the resource in question. In the results below we have represented capacity value in three ways: (1) MW, (2) percent of maximum hourly output for the year, and (3) as percent of rated capacity as indicated by the IOU providing the generation data. In the case of wind, the relatively large discrepancy between actual generation and nameplate generation is probably an artifact of the older technology that still exists in some areas in California. We believe that modern and future wind turbine technology will be more reliable than some past technology has been, minimizing this capacity discrepancy. If wind generation were to receive capacity payments, the wind operator would have an incentive to keep the turbines running and in good repair, especially during high load or LOLP events. Although we generally believe that capacity value should be represented as a percentage of nameplate capacity, this depends on having accurate nameplate values. The PG&E nameplate estimates do not match the maximum wind generation as well as those from SCE. Although this is not conclusive, it suggests that caution should be used in interpreting these capacity values. All of the data issues introduced above are discussed in further detail in Section 5.3.

Table 2.3 below shows the capacity value results expressed in terms of annual peak generation. To clarify, to calculate the relative ELCC for this table, the ELCC (in MW) is divided by the maximum hourly generation, for the resource in question, over the year. Table 2.4 illustrates the results relative to rated capacity, where rated capacity is in the denominator.

Table 2.3. Capacity credit analysis results, based on annual peak generation.

Resource	2002			2003			2004		
	ELCC (MW)	Annual peak generation (MW)	Capacity credit (relative ELCC)	ELCC (MW)	Annual peak generation (MW)	Capacity credit (relative ELCC)	ELCC (MW)	Annual peak generation (MW)	Capacity credit (relative ELCC)
Medium Gas			100%			100%			100%
Biomass	417	427	98%	436	446	98%	456	467	98%
Geothermal (north)	151	139	108%	263	241	109%	262	241	109%
Geothermal (south)	382	351	109%	380	349	109%	380	349	109%
Solar	335	407	82%	314	463	68%	299	401	75%
Wind (Northern Cal)	160	489	33%	170	463	37%	205	462	44%
Wind (San Gorgonio)	138	325	42%	89	317	28%	89	332	27%
Wind (Tehachapi)	168	584	29%	191	568	34%	167	571	29%

The capacity results for wind and solar are different than those of the Phase III report. There are a couple of reasons for these differences. The hydro dataset used in this analysis is actual hydro, hourly, for the full year. In the Phase III work we were constrained to work with modeled hydro. Second, the generation aggregates used for the Phase I and Phase III one-year analyses differ somewhat from those in the current work. As discussed in Section 4.4, the renewable generation data used in the current work is from the IOUs and the composition of the aggregates are not exactly identical to those in the CaISO data previously used. There are differences in the wind and solar data. For example, in 2002 the maximum generation for solar is 16% higher than in Phase III. Maximum wind generation exceeds that in Phase III by 12%, 24%, and 13% for the three wind resource regions. For these reasons, comparability to the previous report is difficult.

The biomass and geothermal results are very close to those obtained in the Phase III report. In the earlier work we also calculated geothermal capacity value based on the time series of actual generation. As we pointed out in the earlier reports, a binding steam constraint would lower the capacity value of the geothermal units in the Geysers. However, we were not able to obtain data that would allow us to distinguish the reason why these geothermal units were operating below rated capacity. It is likely that in some cases the units are responding to dispatch instructions; in other cases the steam constraint may be binding. To properly calculate the capacity value of steam-constrained geothermal units we would need accurate measurements of the *possible* generation for each hour of the year based on steam availability.

Table 2.4 shows the results in terms of the reported nameplate capacity. As can be seen in the table, the percentage capacity values are generally lower than in Table 2.3. Solar is the exception, and this is because there may be times that the combined solar/gas generation can exceed the rated capacity. We were unable to obtain detailed data on the solar generation to validate this hypothesis. For the wind resources, we would expect the capacity value to decline when we use reported capacity as the basis of the capacity value, and we believe that using an accurate measure of nameplate capacity is the most appropriate metric.

Table 2.4. Capacity credit analysis results, based on rated capacity reported by the IOUs.

Resource	2002			2003			2004		
	ELCC (MW)	Reported rated capacity (MW)	Capacity credit (relative ELCC)	ELCC (MW)	Reported rated capacity (MW)	Capacity credit (relative ELCC)	ELCC (MW)	Reported rated capacity (MW)	Capacity credit (relative ELCC)
Medium Gas			100%			100%			100%
Biomass	417	427	98%	436	446	98%	456	467	98%
Geothermal (north)	151	139	108%	263	241	109%	262	241	109%
Geothermal (south)	382	351	109%	380	349	109%	380	349	109%
Solar	335	379	88%	314	379	83%	299	379	79%
Wind (Northern Cal)	160	679	24%	170	679	25%	205	680	30%
Wind (San Gorgonio)	138	357	39%	89	362	24%	89	362	25%
Wind (Tehachapi)	168	652	26%	191	659	29%	167	659	25%

To get an idea of the impact that hydro and interchange have on the LOLP profile, we removed them and re-ran the analysis. We show the results in terms of annual peak generation (Table 2.5) and in terms of reported rated capacity (Table 2.6).

Table 2.5. Capacity credit results with hydro and interchange removed; results based on annual peak generation.

Resource	2002			2003			2004		
	ELCC (MW)	Annual peak generation (MW)	Capacity credit (relative ELCC)	ELCC (MW)	Annual peak generation (MW)	Capacity credit (relative ELCC)	ELCC (MW)	Annual peak generation (MW)	Capacity credit (relative ELCC)
Medium Gas			100%			100%			100%
Biomass	417	427	98%	435	446	98%	456	467	98%
Geothermal (north)	151	139	108%	263	241	109%	262	241	109%
Geothermal (south)	383	351	109%	380	349	109%	380	349	109%
Solar	370	407	91%	332	463	72%	310	401	77%
Wind (Northern Cal)	129	489	26%	129	463	28%	179	462	39%
Wind (San Gorgonio)	124	325	38%	69	317	22%	93	332	28%
Wind (Tehachapi)	175	584	30%	167	568	29%	178	571	31%

A comparison of the wind and solar capacity values from Table 2.5 with the Phase III results shows a much closer correspondence. For example, in the Phase III report Altamont (Northern California) had a capacity value of 26% (based on maximum generation), San Gorgonio 31%, and Tehachapi 29%. The obvious outlier is San Gorgonio. The solar capacity value was 88% compared to 91% here. The relatively good correspondence between some of these values may however be spurious, since there are substantial differences in the data sets used in the two analyses. Assuming accurate data, Table 2.6 provides the most accurate assessment of the capacity values that would have occurred in the absence of interchange and hydro.

Table 2.6. Capacity credit results with hydro and interchange removed; results based on rated capacity reported by the IOUs.

Resource	2002			2003			2004		
	ELCC (MW)	Reported rated capacity (MW)	Capacity credit (relative ELCC)	ELCC (MW)	Reported rated capacity (MW)	Capacity credit (relative ELCC)	ELCC (MW)	Reported rated capacity (MW)	Capacity credit (relative ELCC)
Medium Gas			100%			100%			100%
Biomass	417	427	98%	435	446	98%	456	467	98%
Geothermal (north)	151	139	108%	263	241	109%	262	241	109%
Geothermal (south)	383	351	109%	380	349	109%	380	349	109%
Solar	370	379	98%	332	379	88%	310	379	82%
Wind (Northern Cal)	129	679	19%	129	679	19%	179	680	26%
Wind (San Gorgonio)	124	357	35%	69	362	19%	93	362	26%
Wind (Tehachapi)	175	652	27%	167	659	25%	178	659	27%

Based in part on comments received by Solargenix during the Phase I discussions, we calculated the capacity factor for each renewable based on SCE's definition of the peak period: weekdays during the months of June through September (except holidays) between 12:00 p.m. and 6:00 p.m. As an alternative, we also included the month of May.[4] Table 2.7 shows the results of these calculations based on annual peak generation, and Table 2.8 shows the same information based on rated capacity.

Table 2.7. Capacity factor over peak hours based on annual peak generation.

Resource	2002		2003		2004	
	May through September	June through September	May through September	June through September	May through September	June through September
Biomass	88%	93%	82%	87%	85%	90%
Geothermal (north)	91%	91%	94%	95%	90%	90%
Geothermal (south)	88%	87%	88%	88%	88%	89%
Solar	85%	90%	70%	76%	85%	89%
Wind (Northern Cal)	27%	27%	29%	30%	37%	35%
Wind (San Gorgonio)	41%	39%	28%	26%	34%	30%
Wind (Tehachapi)	36%	33%	28%	28%	33%	29%

Table 2.8. Capacity factor over peak hours based on rated capacity reported by the IOUs.

Resource	2002		2003		2004	
	May through September	June through September	May through September	June through September	May through September	June through September
Biomass	88%	93%	82%	87%	85%	90%
Geothermal (north)	91%	91%	94%	95%	90%	90%
Geothermal (south)	88%	87%	88%	88%	88%	89%
Solar	91%	97%	86%	93%	90%	94%
Wind (Northern Cal)	19%	19%	20%	20%	25%	24%
Wind (San Gorgonio)	37%	36%	25%	23%	31%	28%
Wind (Tehachapi)	32%	30%	24%	24%	29%	25%

Table 2.9 collects results from Table 2.4 and Table 2.8. All ELCC values in the table are expressed as a percentage of the rated capacity and the peak capacity factors are all calculated based on the period from June through September and use rated capacity in the denominator. In each case we also calculated the three-year average. We note that we believe that the combined solar/gas units can generate above rated capacity, and that the capacity factor of the geothermal units (absent steam constraint or dispatch instruction to limit output) is approximately 9% more than the reference unit. In addition, our data for biomass does not tell us the reason why the biomass generation runs below capacity. For some of the wind resource areas we have an excellent match between the three-year average ELCC and the three-year average peak capacity factors. Unfortunately this close match does not extend to the Northern California wind resource, which differs by about 5%.

Table 2.9. ELCC compared to peak capacity factors (June through September, weekdays excluding holidays, 12:00 p.m. to 6:00 p.m.) for three years, based on rated capacity reported by the IOUs.

Resource	2002		2003		2004		3-Year Average	
	ELCC (% of rated capacity)	Peak capacity factor	ELCC (% of rated capacity)	Peak capacity factor	ELCC (% of rated capacity)	Peak capacity factor	ELCC (% of rated capacity)	Peak capacity factor
Biomass	98	93	98	87	98	90	98	90
Geothermal (north)	108	91	109	95	109	90	109	92
Geothermal (south)	109	87	109	88	109	89	109	88
Solar	88	97	83	93	79	94	83	95
Wind (Northern Cal)	24	19	25	20	30	24	26	21
Wind (San Gorgonio)	39	36	24	23	25	28	29	29
Wind (Tehachapi)	26	30	29	24	25	25	27	26

Table 2.10, below, summarizes some of the key results as above, but instead uses ELCC values from the runs that exclude hydro and interchange. All ELCC values in the table are expressed as a percentage of the rated capacity and the peak capacity factors are all calculated based on the period from June through September and use rated capacity in the denominator. In this case we have a match between the results for the Northern California wind area and have a 2% difference in San Gorgonio. Because the hydro and interchange data were removed from these ELCC calculations, the ELCC results are not quite as accurate because of the missing resources. However, because of the lack of interchange and hydro, the relationship between load and LOLP is more straightforward.

Table 2.10. ELCC with hydro and interchange excluded compared to peak capacity factors (June through September, weekdays excluding holidays, 12:00 p.m. to 6:00 p.m.) for three years, based on rated capacity reported by the IOUs.

Resource	2002		2003		2004		3-Year Average	
	ELCC (% of rated capacity)	Peak capacity factor	ELCC (% of rated capacity)	Peak capacity factor	ELCC (% of rated capacity)	Peak capacity factor	ELCC (% of rated capacity)	Peak capacity factor
Biomass	98	93	98	87	98	90	98	90
Geothermal (north)	108	91	109	95	109	90	109	92
Geothermal (south)	109	87	109	88	109	89	109	88
Solar	98	97	88	93	82	94	89	95
Wind (Northern Cal)	19	19	19	20	26	24	21	21
Wind (San Gorgonio)	35	36	19	23	26	28	27	29
Wind (Tehachapi)	27	30	25	24	27	25	26	26

2.4.3 DISCUSSION OF RESULTS

During the prior phases of this project, there have been numerous discussions regarding whether nameplate capacity assignments for existing wind resource areas were correct. Although it is useful to measure capacity value in MW, it is difficult to properly interpret the effectiveness of the resource if its rated capacity is unknown. Although this has been an issue in this project, we believe that it will be less of a problem with new wind facilities. If capacity payments are to be made to renewable (or other) generators, the incentive provided by the payment should be an inducement to ensure generator availability. Perhaps even more important is the evolution in wind turbine technology. Modern turbines are quite unlike many older turbines currently installed in California. Combined with taller towers and larger rotors, energy can be generated at lower wind speeds than with older technology. We expect that the capacity credit, however calculated, will be significantly different for modern/future wind turbines. Going forward, we do not believe that large numbers of turbines will be unaccounted for if good engineering and business practices are followed.

With the uncertainties surrounding data quality during this project, it is hard to know the extent to which data inaccuracies influence the results. We have much better confidence in the revised data sets used for this analysis than in the past. Data confidentiality issues have made it difficult to fully assess the results, particularly given the confidential aggregations of renewable generators.

The ELCC for the renewable generators that were calculated for this report indicate the reliability contribution of the renewable technologies that make up a portion of the generator fleet in California. There are many moving parts that are captured by the model as a snapshot. For example, there may be significant synergies between hydro operations and wind/solar. Based on discussions during this project it appears that the hydro system is dispatched independently of the intermittent generation. With improvements in forecasting, especially for wind, it is possible that some incremental reliability can be gained by exploiting these potential synergies.

We ran several alternative scenarios to calculated ELCC. It is clear that hydro and interchange make a difference in the LOLP profile, and therefore on the ELCC of intermittent generators. It is also evident that ELCC results are not necessarily transparent. We found a reasonably good correspondence between ELCC and capacity factors that were calculated over the peak period. Whether to use ELCC or a capacity factor approximation is a policy decision. The overriding factor that would seem to favor ELCC is that it is a rigorous method that explicitly considers risk via the LOLP equation. Any approximation method will fall short. Conversely, a simpler method such as that considered above can come close and our examples showed that over three years, differences in methods may become less important. Simple methods also have the advantage of transparency and ease of reproduction. However, when a simple method such as a capacity factor approach is applied to the more conventional-appearing renewable technologies such as biomass and geothermal, care must be used to separate dispatch response from capability. In any case, our preferred approach would measure capacity value against an accurate assessment of the installed capacity rating.

We urge all parties to endeavor to collect good data and to run the datasets through a rigorous quality assurance program.

3 REGULATION

3.1 Overview

The method for calculating regulation costs was developed by Brendan Kirby et al at Oak Ridge National Laboratory (ORNL). The methodology and its results are described below with background material presented in Appendix B. The regulation analysis methodology has been applied to a variety of other control areas to quantify the ancillary service impacts of loads and intermittent resources. It determines the regulation and load following impacts to the control area. These impacts are the result of fluctuations in aggregate load and/or uncontrolled generation that must be compensated. Once the requirements are quantified, the method then determines the costs incurred in terms of greater amounts of purchased regulating capacity and greater use of the short-term energy markets.

3.1.1 ANCILLARY SERVICES

Terminology associated with ancillary services has not been standardized across the utility industry and this sometimes has led to confusion. It is important to distinguish between the *impacts* imposed upon the power system and the *resources* or *services* the CaISO utilizes to compensate for these impacts. The impacts in the regulation time frame are imposed upon the power network by loads, uncontrolled generators, and transactions. The resources or services that compensate for these impacts are supplied by generators responding to *automatic generation control* (AGC) and the *automated dispatch system* (ADS).

Regulation and load following are intimately related; both continuously balance aggregate load and generation within the control area. The two services differ in the time frame over which they operate with regulation operating minute-to-minute while load following operates over a ten minute or longer time frame. In 1996 the Federal Energy Regulatory Commission (FERC), defined six ancillary services in its Order 888. This order did not discuss load following. Perhaps because of this omission, most utilities and independent system operators (ISOs) do not include load following in their tariffs. The absence of this service required some ISOs to acquire much more regulation than they otherwise would need. Perhaps because of these problems, FERC, in its notice on regional transmission organizations (RTOs), proposed to require that RTOs operate real-time balancing markets.[5] The responsive resources for these supplemental energy markets are generators that can change output every ten minutes as needed to follow load.

The CaISO obtains responsive resources to achieve the required real-time balancing of generation and load from the hourly regulation markets and the short-term energy markets. The alignment between the impacts that the CaISO must meet and the services it procures to meet those impacts is not perfect. Resources procured through the regulation markets, for example, could be used to provide load following, accommodate energy imbalance, or even supply base energy if there were no other alternatives. Load following itself is not a service which the CaISO procures directly.

The CaISO meets its load following needs through short-term energy transactions, including both AGC generators and the supplemental energy market. Load following results are discussed in Section 4.

3.1.2 DEFINITION OF REGULATION AND LOAD FOLLOWING

Loads within a control area can be decomposed into three components: base energy, load following, and regulation, as shown for a hypothetical weekday morning in Figure 3.1. Starting at a base energy of 3566 MW, the smooth load following ramp (blue) is shown rising to 4035 MW. Regulation (red) consists of the rapid fluctuations in load around the underlying trend, shown here on an expanded scale to the right with a ±55 MW range. Combined, the three elements serve a total load (green) that ranges from 3539 MW to 4079 MW during the 3 hours depicted.

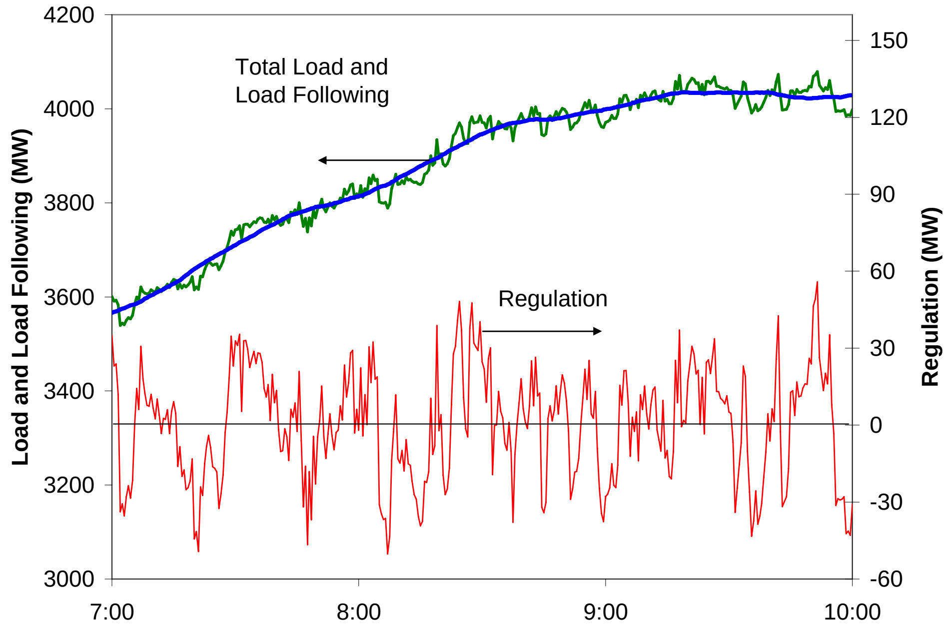

Figure 3.1 Decomposition of hypothetical weekday morning load.

The system responses to the second and third components are called load following and regulation. These two services ensure that, under normal operating conditions, a control area is able to balance generation to load. The two services are briefly defined[6,7,8] as follows:

- *Regulation* is the use of online generating units that are equipped with automatic generation control (AGC) and that can change output quickly (MW/minute) to track the moment-to-moment fluctuations in customer loads and to correct for the unintended fluctuations in generation. In so doing, regulation helps to maintain interconnection frequency, manage differences between actual and scheduled power flows between control areas, and match generation to load within the control area. This service can be provided by any appropriately equipped generator that is connected to the grid and electrically close enough to the local control area that physical and economic transmission limitations do not prevent the importation of this power.

- *Load following* is the use of online generation equipment to track the intra- and inter-hour changes in customer loads. Load following differs from regulation in three important respects. First, it occurs over longer time intervals than does regulation, 10 minutes or more rather than minute to minute. Second, the load-following patterns of individual customers can be highly correlated with each other, whereas the regulation patterns are largely uncorrelated. Third, load-following changes are often predictable (e.g., because of the weather dependence of many loads) and have similar day-to-day patterns.

There is no hard-and-fast rule to define the temporal boundary between regulation and load following. If the time chosen for the split is too short (e.g., five minutes), too much of the fluctuations will appear as load following and not enough as regulation. If the boundary is too long (e.g., 60 minutes), too much of the fluctuations will show up as regulation and not enough as load following. But in each case, the total is unchanged and is captured by one or the other of these two services. A 15-minute rolling average is recommended here to separate regulation from load following. The rolling average for each 1-minute interval should be calculated as the mean value of the seven earlier values of the variable, the current value, and the subsequent seven values. For load:

$$\textit{Load Following}_t = \textit{Load}_{estimated\text{-}t} = \textit{mean}\ (L_{t-7}, L_{t-6}, \ldots, L_t, L_{t+1}, \ldots, L_{t+7}) \qquad \textit{Equation 3.1}$$

$$\textit{Regulation}_t = \textit{Load}_t - \textit{Load}_{estimated\text{-}t} \qquad \textit{Equation 3.2}$$

This method is somewhat arbitrary and imperfect. It is arbitrary in that the time-averaging period (15 minutes as recommended here) and the temporal aggregation of raw data (1 minute) cannot be predetermined. In principle, the control-area characteristics (dynamics of generation and load and the short-term energy market interval) should determine these two factors.[9] The 15-minute rolling average is recommended because it provides good temporal segregation and captures the characteristics of California's supplemental energy market.

In practice, system operators cannot know future values of load. They generally produce short-term forecasts of these values to aid in generation-dispatch decisions. While aggregate load forecasts are typically well developed, and a short-term energy market now operates in California, short-term forecast methodologies for non-dispatchable conventional and renewable generators are not. The rolling average has proven to be a reasonable analytical substitute in studying other control areas. The rolling average, like the system operator through the use of the short-term energy market, is constantly moving the regulating units back to the center of their operating range. If consistent, robust short-term forecasts are available and verified for all of the renewable generation technologies, this analysis can be performed without the use of a rolling average.

The use of the rolling average rather than the short term forecasts can impact the allocation of variability between the regulation and load following services slightly. Significantly, the method assures that total variability is captured in one or the other service and that there is no double counting.

The distinctions between regulation and load following are discussed further in Section 4.1

3.2 Regulation Analysis Methodology

The regulation analysis methodology quantifies the regulation impacts of loads and generating resources within a control area. These impacts are the result of fluctuations in aggregate load and/or uncontrolled generation that must be compensated. Once the requirements are quantified, the method then determines the costs incurred in terms of greater amounts of purchased regulating capacity.

The regulation requirement of the entire system is first determined by taking the standard deviation of the 1 minute regulation values (applying Equation 3.2) for the total system. This is done hourly because the regulation market clears hourly. It is then possible to calculate individual contributions to that total requirement. Regulation aggregation is nonlinear; there are strong aggregation benefits. It takes much less regulation effort to compensate for the total aggregation than it would take if each load or generator compensated for its regulation impact individually. An allocation method should:

- Recognize positive and negative correlations
- Be independent of sub-aggregations
- Be independent of the order in which loads or resources are added to the system
- Allow dis-aggregation of as many or few components as desired

The method presented here, and described more fully in Appendix B, meets these criteria. It was developed to analyze the impacts of nonconforming loads on power system regulation and works equally well when applied to non-dispatchable or uncontrolled generators. The allocation method does not require knowledge of each individual's contribution to the overall requirement. Specific individual contributions can be calculated based upon the total requirement and the individual's performance. Because regulation is composed of short, minute-to-minute fluctuations, the regulation component of each individual is often largely uncorrelated with those of other individuals. If each individual's fluctuations (represented by the standard deviation, σ_i) is completely independent of the remainder of the system, the total regulation requirement (σ_T) would equal:

$$\sigma_T = \sqrt{\sum \sigma_i^2} \qquad \textit{Equation 3.3}$$

where i refers to an individual and T is the system total

For the case of uncorrelated contributions, the share of regulation assigned to each individual is:

$$Share_i = \left(\frac{\sigma_i}{\sigma_T}\right)^2 \qquad \textit{Equation 3.4}$$

The more general allocation method, presented in Equation 3.5, accommodates any degree of correlation and any number of individuals. This allocation method is more complex but no more data-intensive than the previous method. This method yields results that are independent of any sub-aggregations. In other words, the assignment of regulation to generator (or load) g_i is not dependent on whether g_i is billed for regulation independently of other non-AGC generators (or loads) or as part of a group. In addition, the allocation method rewards (pays) generators (or loads) that reduce the total regulation impact.

$$Share_i = \frac{\sigma_T^2 + \sigma_i^2 - \sigma_{T-i}^2}{2\sigma_T} \qquad \textit{Equation 3.5}$$

The general allocation method (Equation 3.5) is recommended for analysis of the impacts of various individual renewable generators on the overall system's regulation requirements.

Calculated hourly regulation requirements are compared with actual hourly regulation purchases by the CaISO and hourly regulation self-provided by scheduling coordinators. Typically, three to five standard deviations of regulating reserves are carried to assure adequate CPS (Control Performance Standards) performance (see Section 4.1 and Appendix A for a discussion of CPS). Total regulation requirements are then allocated back to individuals. Hourly regulation costs are used to allocate the cost of regulation back to individuals. All of the CaISO's regulation requirements are allocated based upon the short-term variability impacts of the loads and renewable generators.

3.3 Data Requirements

Studying regulation requires one-minute, synchronized, integrated-energy, time series data for total control area load and the individual renewable resources of interest.

At a minimum, the data list must include time series data for:

- Total load
- Each renewable generator of interest

Experience has shown that it is also wise to perform an energy balance around the control area to assure data integrity. This requires 1-minute data for total generation, net actual imports/exports, net scheduled imports/exports, system frequency (and the frequency bias), and ACE. The data list should include one minute, synchronized, integrated-energy, time series data for:

- Total generation
- Net actual imports/exports
- Net scheduled imports/exports
- Area control error (ACE)
- Frequency (and frequency bias) – often provided as a deviation from scheduled frequency

Regulation analysis requires only one system data element plus one for each renewable generator of interest, each minute. Verifying data integrity requires an additional five system data elements each minute.

The CaISO runs hourly markets for regulation up and regulation down. Price and quantity data from these markets are used to determine practical quantities and costs of procured regulating resources. Scheduling coordinators are also allowed to self-provide regulation. The amount of self-provided regulation must be added to the amount of purchased regulation to obtain the total regulation amount. There is no price associated with self-provided regulation so the market price of the purchased regulation for the same hour is used to calculate the total dollar value of regulation for each hour.

- Hourly regulation-up price
- Hourly regulation-down price
- Hourly MW of regulation-up procured (hour ahead and real-time)
- Hourly MW of regulation-down procured (hour ahead and real-time)
- Hourly MW of regulation-up self-provided
- Hourly MW of regulation-down self-provided

3.4 Step-by-Step Regulation Analysis Methodology

The following is a step-by-step listing of the regulation analysis. Inputs are explicitly listed as they are newly introduced into the calculations.

1. Verify data consistency by looking at total system inflows, outflows, generation, and load.

$$ACE(t) = [NI_A(t) - NI_S(t)] - 10\beta[(F_A(t) - F_S(t)] - I_{ME}(t) \quad \textit{Equation 3.6}$$

$$NI_A(t) = G(t) - L(t) \quad \textit{Equation 3.7}$$

Table 3.1. Regulation inputs/outputs: Verify data consistency.

Inputs

	Data description		Units	Sampling rate
a.	L	total actual system load	MW	1 minute
b.	G	total actual system generation	MW	1 minute
c.	F_A	actual system frequency	Hz	1 minute
d.	F_S	scheduled system frequency	Hz	1 minute
e.	ACE	area control error	MW	1 minute
f.	NI_A	actual net tie flows	MW	1 minute
g.	NI_S	scheduled net tie flows	MW	1 minute
h.	β	control area frequency bias	$MW/0.1\,Hz$	1 minute

2. Calculate the total (net) system compensation requirement for each time step by subtracting the measured generators from the total actual system load.

$$\begin{aligned} L_T &= L - \sum g_i \\ &= L - g_B - g_G - g_S - g_W - g_C \end{aligned} \quad \textit{Equation 3.8}$$

Table 3.2. Regulation inputs/outputs: Calculate total system compensation requirement.

Inputs

	Data description		Units	Sampling rate
a.	L	total system load	MW	1 minute
b.	g_B	biomass generation	MW	1 minute
c.	g_G	geothermal generation	MW	1 minute
d.	g_S	solar generation	MW	1 minute
e.	g_W	wind generation	MW	1 minute
f.	g_C	sample conventional generation	MW	1 minute

Outputs

	Data description		Units	Sampling rate
a.	L_T	total system compensation requirement	MW	1 minute

3. Calculate 15 minute rolling average to use as a surrogate for the short term forecast.

$$L_{T,ave}(t) = \overline{L_{T,15}}(t) = \frac{\sum_{x=-7\min}^{7\min} L_T(t+x)}{15} \quad \textit{Equation 3.9}$$

$$L_{ave}(t) = \overline{L_{15}}(t) = \frac{\sum_{x=-7\min}^{7\min} L(t+x)}{15} \quad \textit{Equation 3.10}$$

$$g_{i,ave}(t) = \overline{g_{i,15}}(t) = \frac{\sum_{x=-7\min}^{7\min} g_i(t+x)}{15} \quad \textit{Equation 3.11}$$

Table 3.3. Regulation inputs/outputs: Estimate short term forecast with 15 minute rolling average.

Inputs

	Data description		Units	Sampling rate
a.	L_T	total system compensation requirement	MW	1 minute

Outputs

	Data description		Units	Sampling rate
a.	$L_{T,ave}$	short term forecast of total system compensation	MW	1 minute
b.	L_{ave}	short term load forecast	MW	1 minute
c.	$g_{B,ave}$	short term forecast of biomass generation	MW	1 minute
d.	$g_{G,ave}$	short term forecast of geothermal generation	MW	1 minute
e.	$g_{S,ave}$	short term forecast of solar generation	MW	1 minute
f.	$g_{W,ave}$	short term forecast of wind generation	MW	1 minute
g.	$g_{C,ave}$	short term forecast of sample conventional generation	MW	1 minute

4. Calculate the raw regulation component by subtracting the short term forecast from the actual data.

$$r_T(t) = L_T(t) - L_{T,ave}(t) \qquad \textit{Equation 3.12}$$

$$r_L(t) = L(t) - L_{ave}(t) \qquad \textit{Equation 3.13}$$

$$r_i(t) = g_i(t) - g_{i,ave}(t) \qquad \textit{Equation 3.14}$$

Table 3.4. Regulation inputs/outputs: Calculate raw regulation component by subtracting short term forecast.

Outputs

	Data description		Units	Sampling rate
a.	r_T	regulation component of total system compensation requirement	MW	1 minute
b.	r_L	regulation component of total system load	MW	1 minute
c.	r_B	regulation component of biomass generator(s)	MW	1 minute
d.	r_G	regulation component of geothermal generator(s)	MW	1 minute
e.	r_S	regulation component of solar generator(s)	MW	1 minute
f.	r_W	regulation component of wind generator(s)	MW	1 minute
g.	r_C	regulation component of sample non-controlled conventional generator(s)	MW	1 minute

5. Calculate the difference between the regulation component of the resource of interest and the regulation component of the total system compensation requirement. The difference is the total system regulation requirement if the resource of interest was not present.

$$\Delta r_i(t) = r_T(t) - r_i(t) \qquad \textit{Equation 3.15}$$

Table 3.5. Regulation inputs/outputs: Calculate total system regulation less resource of interest.

Outputs

	Data description		Units	Sampling rate
a.	Δr_L	total system regulation without load	MW	1 minute
b.	Δr_B	total system regulation without biomass generator(s)	MW	1 minute
c.	Δr_G	total system regulation without geothermal generator(s)	MW	1 minute
d.	Δr_S	total system regulation without solar generator(s)	MW	1 minute
e.	Δr_W	total system regulation without wind generator(s)	MW	1 minute
f.	Δr_C	total system regulation without sample conventional generator(s)	MW	1 minute

6. Calculate the hourly standard deviation of the regulation values determined in the previous two steps.

$$\sigma_T(t) = \underset{x=0 \rightarrow 59\,\min}{\sigma}\left(r_T(t+x)\right) \qquad \textit{Equation 3.16}$$

$$\sigma_i(t) = \underset{x=0 \rightarrow 59\,\min}{\sigma}\left(r_i(t+x)\right) \qquad \textit{Equation 3.17}$$

$$\sigma_{T-i}(t) = \underset{x=0 \rightarrow 59\,\min}{\sigma}\left(\Delta r_i(t+x)\right) \qquad \textit{Equation 3.18}$$

Table 3.6. Regulation inputs/outputs: Calculate statistical metrics.

Outputs

	Data description		Units	Sampling rate
a.	σ_T	standard deviation of regulation component of total system requirement	MW	1 hour
b.	σ_L	standard deviation of regulation component of total system load	MW	1 hour
c.	σ_B	standard deviation of regulation component of biomass generator(s)	MW	1 hour
d.	σ_G	standard deviation of regulation component of geothermal generator(s)	MW	1 hour
e.	σ_S	standard deviation of regulation component of solar generator(s)	MW	1 hour
f.	σ_W	standard deviation of regulation component of wind generator(s)	MW	1 hour
g.	σ_C	standard deviation of regulation component of sample non-controlled conventional	MW	1 hour

		generator(s)		
h.	σ_{T-L}	standard deviation of regulation of system without load	MW	1 hour
i.	σ_{T-B}	standard deviation of regulation of system without biomass generator(s)	MW	1 hour
j.	σ_{T-G}	standard deviation of regulation of system without geothermal generator(s)	MW	1 hour
k.	σ_{T-S}	standard deviation of regulation of system without solar generator(s)	MW	1 hour
l.	σ_{T-W}	standard deviation of regulation of system without wind generator(s)	MW	1 hour
m.	σ_{T-C}	standard deviation of regulation of system without sample conventional generator(s)	MW	1 hour

7. Allocate the regulation standard deviation share to load and each resource of interest.

$$\hat{R}_i(t) = Share_i(t) = \frac{\sigma_T^2(t) + \sigma_i^2(t) - \sigma_{T-i}^2(t)}{2\sigma_T(t)}$$ *Equation 3.19*

Table 3.7. Regulation inputs/outputs: Allocate regulation share for each generator.

Outputs

	Data description		Units	Sampling rate
a.	$\hat{R}_L$	regulation standard deviation share of total system load	MW	1 hour
b.	$\hat{R}_B$	regulation standard deviation share of biomass generation	MW	1 hour
c.	$\hat{R}_G$	regulation standard deviation share of geothermal generation	MW	1 hour
d.	$\hat{R}_S$	regulation standard deviation share of solar thermal generation	MW	1 hour
e.	$\hat{R}_W$	regulation standard deviation share of wind generation	MW	1 hour
f.	$\hat{R}_c$	regulation standard deviation share of sample conventional generation	MW	1 hour

8. Determine the actual regulation requirement of the total system load and each resource of interest. We assume that the CaISO is currently purchasing the correct amount of regulation and appropriately controlling the system to achieve a good balance of cost and reliability performance. We allocated the amount and cost of regulation to the aggregated loads and selected renewable generators.

$$R_i(t) = \frac{\hat{R}_i(t) R_{actual}(t)}{\sigma_T(t)}$$ *Equation 3.20*

Table 3.8. Regulation inputs/outputs: Calculate actual regulation share for each generator type.

Inputs

	Data description		Units	Sampling rate
a.	R_{actual}	actual regulation (purchased and self provided, up and down) market data	MW	1 hour

Outputs

	Data description		Units	Sampling rate
a.	R_L	regulation requirement of total system load	MW	1 hour
b.	R_B	regulation requirement of biomass generator(s)	MW	1 hour
c.	R_G	regulation requirement of geothermal generator(s)	MW	1 hour
d.	R_S	regulation requirement of solar generator(s)	MW	1 hour
e.	R_W	regulation requirement of wind generator(s)	MW	1 hour
f.	R_C	regulation requirement of sample conventional generator(s)	MW	1 hour

9. Calculate actual hourly regulation cost by multiplying the actual regulation requirement by hourly regulation cost. Calculate the change in cost that results from each renewable generator.

$$COST_R(t) = R_i(t) \cdot RATE_R(t)$$ *Equation 3.21*

Table 3.9. Regulation inputs/outputs: Calculate actual regulation cost for each generator type.

Inputs

	Data description		Units	Sampling rate
a.	$RATE_R$	actual regulation rate (up an down) market data	$/MW-hr	1 hour

Outputs

	Data description		Units	Sampling rate
a.	$COST_{R,L}$	regulation cost of total system load	$	1 hour
b.	$COST_{R,B}$	regulation cost of biomass generator(s)	$	1 hour
c.	$COST_{R,G}$	regulation cost of geothermal generator(s)	$	1 hour
d.	$COST_{R,S}$	regulation cost of solar generator(s)	$	1 hour
e.	$COST_{R,W}$	regulation cost of wind generator(s)	$	1 hour
f.	$COST_{R,C}$	regulation cost of sample conventional generator(s)	$	1 hour

3.4.1 ANALYSIS CHANGES FROM PHASE I AND PHASE III

In the spring of 2005, an independent review* of the Phase I report revealed that the calculation of the total system compensation requirement did not include the renewable generators' variability along with the total load variability. Only the total load variability was included. The methodology implementation description above, specifically Equation 3.8, now explicitly includes the individual generators as well as the load.

Later, a one minute data misalignment was discovered in the wind data for San Gorgonio used in the Phase I analysis. The misalignment only affected the regulation results because its effect is suppressed by the hourly and ten minute averaging used by the capacity credit and load following calculations. A revised set of results for the Phase I regulation analysis is presented below. This includes the complete calculation of the total system compensation requirement and synchronized data for San Gorgonio.

* The independent review was performed by Matthew Barmack of Analysis Group, Inc. When he could not duplicate the Phase I regulation results, we investigated further and found the omission in the total system compensation calculation. We are grateful to Matthew

Table 3.10. Original and corrected results of the Phase I (one year, 2002) regulation analysis. Negative values are costs to the system.

Resource	Regulation Cost ($/MWh or mills/kWh)	
	Original	Corrected
Total System	-0.42	-0.44
Total Load	-0.42	-0.41
Medium Gas	0.08	-0.28
Biomass	0.00	-0.09
Geothermal	-0.10	-0.17
Solar	0.04	-0.47
Wind (Altamont)	0.00	-0.22
Wind (San Gorgonio)	-0.46	-0.08
Wind (Tehachapi)	-0.17	-0.53
Wind (Total)	-0.17	-0.33

The results for the total system and for load remain approximately the same because the load represents the majority of variability in the entire system. However, because the variability of the individual generators was not originally included in the total system regulation requirement, the amount of variability allocated to each generator was understated. The decrease in San Gorgonio is not a result of including its variability in the total regulation requirement, but because of the correction of the one minute misalignment in its generation data (a calculation with the original misaligned data indeed results in a cost increase). The cost for San Gorgonio is several times lower than the other wind regions. This may be an anomaly, as shown in the multi-year results for San Gorgonio, below. The results are discussed further along with the multi-year analysis results in the following section.

The datasets used in the multi-year analysis vary somewhat from the datasets used in the Phase I one year analysis. The CaISO multi-year dataset has expanded aggregates in an attempt to better represent the generators being studied. However, the multi-year dataset exhibited new types of errors. To address these errors, the multi-year dataset was reviewed and checked for errors using data from PG&E and SCE as bases of comparison. The multi-year analysis replaced the Altamont aggregate with an aggregate including plants from Altamont, Solano, and Pacheco; this was necessary to more closely match the corresponding PG&E data aggregate that it was compared against. Because of gaps in the 2002 biomass and solar data, the 2002 biomass and solar regulation analyses were run normally, but the runs for the other generation aggregates excluded biomass and solar from their calculation of the total system compensation requirement. This was considered a reasonable approximation because results from the 2002 one year analysis are available for comparison. All of the data issues are detailed in Section 5.3.

3.5 Multi-Year Analysis Results and Discussion

The methodology described above was applied to the CaISO multi-year dataset. The results of the multi-year analysis appear below.

Table 3.11. Results of regulation analysis of multi-year dataset. Negative values are a cost.

Resource	Regulation Cost ($/MWh or mills/kWh)		
	2002	2003	2004
Total System	-0.42	-0.47	-0.39
Total Load	-0.41	-0.46	-0.36
Biomass	-0.09	-0.13	-0.12
Geothermal	-0.11	-0.03	-0.02
Solar	-0.44	-0.47	-0.37
Wind (Northern California)	-0.24	-0.40	-0.33
Wind (San Gorgonio)	-0.09	-0.43	-0.58
Wind (Tehachapi)	-0.57	-0.70	-0.56
Wind (Total)	-0.36	-0.53	-0.47

Note: Use caution when applying $/MWh as a regulation cost metric.

Using $/MWh as a metric for regulation is both useful and dangerous. It is useful because what we really want to know is how much this ancillary service (something we are forced to buy but don't really want) adds to the cost of electricity (something that does useful work for us and we do want to purchase). In that sense a metric that is in the same units ($/MWh) as the commodity we are purchasing is very useful. It is dangerous because the amount of regulation required and the price have almost nothing to do with the amount of energy consumed or produced. The amount of regulation depends upon the short-term volatility of the generation or load, not the energy consumption or production. Use $/MWh in reference to regulation with great caution.

The 2002 results from the multi-year analysis and the one year analysis (Table 3.10) match well. There is some minor variation, but this is expected as the composition of the generation aggregates are not exactly identical. The effect of the 2002 biomass and solar data gaps in the multi-year dataset was negligible. Indeed, the biomass values match exactly and the solar values are very close.

In general, regulation costs increased slightly from 2002 to 2003 and then fell again in 2004, although not to previous levels. The calculated regulation purchase amount and costs are scaled from actual regulation commitment and purchase data from the CaISO OASIS database, which is shown below in Table 3.12.

Table 3.12. Actual regulation amounts committed in the CaISO control area, 2002-2004.

	2002	2003	2004
Regulation up, self provided (MW-hr)	1,855,270	1,769,493	1,972,175
Regulation down, self provided (MW-hr)	2,078,057	1,797,975	2,073,533
Regulation up, procured (MW-hr)	1,659,438	1,116,009	1,109,265
Regulation down, procured (MW-hr)	1,627,342	1,488,440	1,255,973
Total regulation (MW-hr)	7,220,107	6,171,916	6,410,947
Total value ($)	98,270,561	109,357,025	88,141,708
Average regulation price ($/MW-hr)	13.61	17.72	13.75

In the table above, note that MW-hr is the commitment of one MW of capacity for one hour and is not the same as MWh, a unit of energy. Also, as stated above, there is no price associated with self-provided regulation so the market price of the purchased regulation for the same hour is used to calculate the total dollar value of regulation for each hour.

Between 2002 and 2003, the actual amount of regulation committed over the entire CaISO control area decreased by 15%. However, the average price increased by 30%, resulting in a net increase in cost of 11%. From 2003 to 2004, the amount of regulation committed stayed approximately the same with a 4% increase. The price returned to 2002 levels resulting in a net cost decrease of 19% between 2003 and 2004.

The calculated regulation costs for the total system requirement and total load follow this pattern closely. In all three years, the regulation costs of the total load are very close to that of the total system requirement, a result of the sheer size of the load. The results could have been different only if one or more of the other studied resources had a dramatic regulation impact. A single large arc furnace, for example, would have sufficient impact to alter the cost of regulation for the rest of the load. None of the resources studied have that sort of regulation impact. In fact, the generating resources studied have quite minor impacts on total system regulation requirements.

Ignoring the outlying low value of San Gorgonio in 2002 for now, the regulation costs of the wind aggregates range from $0.24/MWh to $0.70/MWh. Not unexpectedly the wind plants impose a small regulation burden on the power system within the same order of magnitude as load when evaluated on a per MWh basis. This was expected because there is no apparent mechanism that would tie the wind plant performance to the power system's needs in the regulation time frame. The regulation burden is low because there is no mechanism that ties wind plant fluctuations to aggregate load fluctuations in a compounding way either. Wind and load minute-to-minute fluctuations appear to be uncorrelated. Hence they greatly benefit from aggregation.

The variation in regulation costs across the three wind regions may be a result of geography, technology, and turbine numbers. The Northern California wind aggregate, for example, has lower costs all three years than the other two regions (again, ignoring San Gorgonio in 2002), possibly because it is composed of the largest numbers of turbines[10].

The inter-annual changes in regulation costs for Tehachapi follow the overall trend of actual regulation commitment in the CaISO control area. The Northern California wind aggregate does too, but to a lesser extent between 2003 and 2004 when the cost increased 67%. San Gorgonio is unique among all the resources studied, showing a 378% jump between 2002 and 2003 and then further increase instead of a decline between 2003 and 2004. The $.09/MWh value for 2002 is significantly lower than any of the other annual wind regulation results. San Gorgonio's individual variability, as defined in Equation 3.17, is not significantly lower in 2002 than 2003. There are also no known mechanisms that would correlate (or not correlate) its fluctuations in the regulation time frame to the rest of the system any differently in 2002 than in any other year. The 2002 value therefore remains anomalous. It was confirmed with the results from the analysis of the 2002 one year dataset, but it may be possible that there are underlying, undetected issues with the 2002 San Gorgonio data in both the one year and multi-year datasets. The 2003 and 2004 results are more consistent with the results of the other regional wind aggregates.

The geothermal aggregate shows a small regulation burden in 2002 which drops off in the later years. A plant with steady output would be expected to impose little or no regulation burden, as seen in 2003 and 2004. However, in 2002, regulation costs may have increased because of the block scheduling from January to May, as shown in Figure 5.3 and discussed in Section 5.2.4.

The biomass aggregate has a consistently low regulation cost with inter-annual changes tracking the changes in actual purchases. Solar showed higher regulation costs, which is consistent with minute-to-minute fluctuations evident in its generation data. The variability in the solar data was greater than what might be expected from a pure solar installation and may be an effect of auxiliary gas generators as they maneuver to meet peaks and follow price signals. The partially controllable nature of the solar plants may have also minimized the cost increase between 2002 and 2003.

Overall, the regulation analysis results are reasonable. Because (1) inter-annual variations exhibited by some resources were disproportionate to changes in actual purchases amounts, (2) large amounts of new capacity will be installed in the future, and (3) technology and operation changes may have a significant effect, the continued understanding of regulation impacts and costs would benefit from more analysis over future years. Analysis with the methodology as described remains straightforward, given the availability of sufficient quality data.

4 LOAD FOLLOWING

In this section we will focus on the renewable resource impacts in the load following time frame, which generally encompasses periods ranging from ten minutes up to a few hours.

4.1 Overview

Load and generation must be continuously balanced on a nearly instantaneous basis in an electric power system. This is one of the characteristics that makes supplying electricity different from providing any other public good such as natural gas, water, telephone service, or air traffic control. It is a physical requirement that does not depend on the market structure. How load and generation are balanced does depend, in part, on the structure of the electricity markets. One benefit of interconnecting multiple control areas is that balancing load and generation within a single control area does not have to be perfect. The North American Electric Reliability Council (NERC) has established rules governing how well each control area must balance load and generation. Control Performance Standards 1 and 2 (CPS1 & CPS2, discussed further in Appendix A) establish statistical limits on how well each control area must balance minute-to-minute fluctuations. Inadvertent interchange accounts track longer term differences. In all cases the total system remains in balance (otherwise blackouts occur). When one control area fails to balance its load with its generation, generation in another control area provides the balance.

The balancing of aggregate load with aggregate generation is accomplished through several services that are distinguished by the time frame over which they operate. As discussed above in Section 3.1.2, regulation and load following (which, in competitive spot markets such as in California, is provided by the intra-hour workings of the real-time energy market) are the two services required to continuously balance generation and load under normal conditions[9]. There is no hard-and-fast rule to define the temporal boundary between regulation and load following. In the PJM region, New York, New England, and Ontario, load following is defined as the 5 minute ramping capability of a generator. In Texas it is a 15 minute service, and in Alberta and California it is a 10 minute service.

Interestingly, control area operators do not need to specifically procure load following; it is obtained from the short-term energy market with generators responding to real-time energy prices. In the CaISO control area, this is known as the supplemental energy market. Regulation, however, requires faster response than can be obtained from units responding to market signals alone. Instead, generators (and potentially storage and/or responsive load) offer capacity that can be controlled by the system operator's AGC system to balance the power system.

Control areas are not able and not required to perfectly match generation and load. NERC has established the Control Performance Standard (CPS) to determine the amount of imbalance that is permissible for reliability purposes. CPS1 measures the relationship between the control area's area control error (ACE, see Appendix A) and the interconnection frequency on a 1 minute average basis. CPS1 values can be either "good" or "bad." When frequency is above its reference value, undergeneration benefits the interconnection by lowering frequency and leads to a good CPS1 value. Overgeneration at such times, however, would further increase frequency and lead to a bad CPS1 value. CPS1, although recorded every minute, is evaluated and reported on an annual basis. NERC sets minimum CPS1 requirements that each control area must exceed each year.

CPS2, a monthly performance standard, sets control-area-specific limits on the maximum average ACE for every 10 minute period. Control areas are permitted to exceed the CPS2 limit no more than 10% of the time. This 90% requirement means that a control area can have no more than 14.4 CPS2 violations per day, on average, during any month.

4.2 Methodology Description

Integration of large amounts of renewable generators could potentially increase errors between scheduled and actual generation. Increases in scheduling error combined with the existing error in load forecasting could change the composition or size of the "generator stack" which responds to load following needs. If such a distortion of the generator stack occurred it could shift the market to marginal generators, whose costs are higher. That could increase the price of energy across the market and thus create implicit costs which were imposed on the entire system by the renewable generators.

The analysis focused on the potential impacts to the generator stack caused by scheduling error. The methodology looks at the impact of renewable generators on the total system scheduling error. If renewable generators create systematic errors that significantly increase the need for generation resources, then they could have a material effect on the composition of the generator stack or the ex-post price for energy.

The analysis methodology first determines system forecasting and scheduling errors for a benchmark case without renewable generators. CaISO prepares hour ahead forecasts of its generation requirements, which represent its best estimate of actual system load. The scheduling coordinators provide schedules for generation which are designed to economically meet the forecasted needs. The scheduling coordinators typically schedule significantly less generation than is needed during peak demand periods and rely upon the hour ahead market to provide the balance. The difference between the forecasted load and the scheduled load is defined as the scheduling bias. Forecast and scheduling errors in the benchmark case provide an indication of the variability inherent in operating the utility grid and are important because they define the normal range of errors without renewable generation impacts.

The scheduling errors for each renewable generator under study are then calculated. The difference between the actual and forecasted load is the load forecasting error. Worst case scheduling was used to estimate the impacts of the renewable generators. Bids for the hour ahead market are due 150 minutes prior to each market cycle. The scheduled output for the hour ahead market was defined by a simple persistence model, assuming that output 150 minutes in the future would be equal to output at the present time. For solar generators it was assumed that scheduled output was equal to what it had been on the previous day at the same time period.

The total system error including the renewable resources was calculated by combining the system forecast error (without renewables) with the additional scheduling error produced by the renewable resources. The forecasting error including renewable generators was then compared against the benchmark case and reviewed to identify significant differences. The goal of this analysis was to determine if the renewable resources significantly changed the total system error, thereby potentially modifying the generator bid stack.

4.2.1 STEP-BY-STEP LOAD FOLLOWING ANALYSIS METHODOLOGY

The following is a step-by-step listing of the analysis methodology for studying the impact of forecasting and scheduling errors. Inputs are explicitly listed as they are newly introduced into the calculations.

1. Calculate the system forecasting error, defined as the difference between the hour ahead forecast prepared by CaISO and the actual system load. (8760 hourly values.)

$$e_{Forecast}(t) = L_{HA\_Forecast}(t) - L_{Actual}(t) \qquad \textit{Equation 4.1}$$

Table 4.1. Load following inputs/outputs: Calculate load forecasting error.

Inputs

	Data description		Units	Sampling rate
a.	$L_{HA\_Forecast}$	Hour ahead load forecast	MW	1 hour
b.	L_{Actual}	Actual load	MW	1 hour

Outputs

	Data description		Units	Sampling rate
a.	$e_{Forecast}$	Load forecasting error	MW	1 hour

2. Calculate the system scheduling error, defined as the difference between the hour ahead schedule provided by the scheduling coordinators and the actual system load. (8760 hourly values.)

$$e_{Schedule}(t) = L_{HA\_Schedule}(t) - L_{Actual}(t) \qquad \textit{Equation 4.2}$$

Table 4.2. Load following inputs/outputs: Calculate system scheduling error.

Inputs

	Data description		Units	Sampling rate
a.	$L_{HA\_Schedule}$	Hour ahead generation schedule	MW	1 hour

Outputs

	Data description		Units	Sampling rate
a.	$e_{Schedule}$	Scheduling error	MW	1 hour

3. Calculate the system scheduling bias, defined as the difference between the hour ahead schedule and the hour ahead forecast. (8760 hourly values.)

$$e_{Bias}(t) = L_{HA\_Schedule}(t) - L_{HA\_Forecast}(t) \qquad \textit{Equation 4.3}$$

Table 4.3. Calculate system scheduling bias.

Outputs

	Data description		Units	Sampling rate
a.	e_{Bias}	Scheduling bias	MW	1 hour

4. Calculate the hour ahead schedule of the generators of interest assuming a "worst-case," simple persistence model. The hour ahead schedule is prepared 150 minutes ahead of time. The persistence model assumes that the generation at time t is equal to the output 150 minutes ago at time t-150. With hourly data, generation data for t-150 (2.5 hours ago) is unavailable, so an average of generation at t-120 and t-180 (two and three hours ago) is used instead. For solar, the model assumed that generation at a given time is equal to the generation at the same time the previous day.

$$\begin{aligned} & g_{i,HA}(t) = g_i(t-150) \\ \text{except for solar: } & g_{S,HA}(t) = g_S(t-1440) \\ \text{where: } & g_i \text{ is actual generation, and} \\ & g_{i,HA} \text{ is the hour ahead schedule} \end{aligned} \qquad \textit{Equation 4.4}$$

Table 4.4. Load following inputs/outputs: Calculate hour ahead schedule of generation resources.

Inputs

	Data description		Units	Sampling rate
a.	g_B	biomass generation	MW	1 hour
b.	g_G	geothermal generation	MW	1 hour
c.	g_S	solar generation	MW	1 hour
d.	g_W	wind generation	MW	1 hour

Outputs

	Data description		Units	Sampling rate
a.	$g_{B,HA}$	hour ahead schedule of biomass generation	MW	1 hour
b.	$g_{G,HA}$	hour ahead schedule of geothermal generation	MW	1 hour
c.	$g_{S,HA}$	hour ahead schedule of solar generation	MW	1 hour
d.	$g_{W,HA}$	hour ahead schedule of wind generation	MW	1 hour

5. Calculate the scheduling error of the generation resources. The scheduling error is defined to be the difference between the resource's load following generation component and its hour ahead schedule. The load following and regulation components of generation can be decomposed as discussed in Section 3.1.2. (8760 hourly values.)

$$e_i(t) = g_{i,lf}(t) - g_{i,HA}(t) \qquad \textit{Equation 4.5}$$

Table 4.5. Load following inputs/outputs: Calculate the resource scheduling error.

Inputs

	Data description		Units	Sampling rate
a.	$g_{B,lf}$	load following generation component of biomass generator(s)	MW	1 hour
b.	$g_{G,lf}$	load following generation component of geothermal generator(s)	MW	1 hour
c.	$g_{S,lf}$	load following generation component of solar generator(s)	MW	1 hour
d.	$g_{W,lf}$	load following generation component of wind generator(s)	MW	1 hour

Outputs

	Data description		Units	Sampling rate
b.	e_B	scheduling error for biomass generator(s)	MW	1 hour
c.	e_G	scheduling error for geothermal generator(s)	MW	1 hour
d.	e_S	scheduling error for solar generator(s)	MW	1 hour
e.	e_W	scheduling error for of wind generator(s)	MW	1 hour
f.	e_C	scheduling error for sample conventional generator(s)	MW	1 hour

4.2.2 ANALYSIS CHANGES FROM PHASE I AND PHASE III

The Phase I load following analysis used minute-to-minute generation data averaged over fifteen minute and one hour intervals. Because of data quality concerns with the one minute data in the multi-year dataset, high quality hourly data from the IOUs was used instead in the multi-year analysis. In Step 5, above, the hourly generation values are used directly as the load following generation components. This is not expected to affect the results.

4.3 Multi-Year Analysis Results and Discussion

The load forecasts prepared by CalSO provide the best estimate of the upcoming system load conditions. Figure 4.1 presents a graphical comparison of the hour ahead forecast load and the actual load for an example period of several days. Since it is not possible to perfectly predict the load in the hour ahead time frame, there will always be some forecast error.

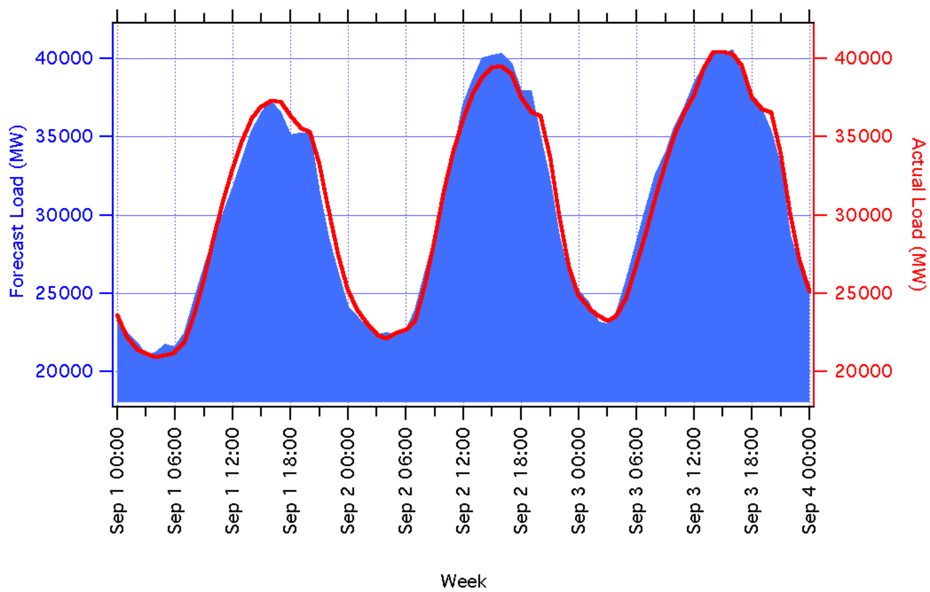

Figure 4.1. Forecast and actual load over a three day sample period.

The load schedule is created by the scheduling coordinators based on forecast information from CalSO and conditions in the energy markets. The hour ahead schedule as compared to the actual load is presented in Figure 4.2 for several example days in September. During peak hours the scheduled load is typically well below the actual load with the difference made up by the hour ahead market. This indicates that the hour ahead market can be relied upon for large amounts of power to meet short term needs.

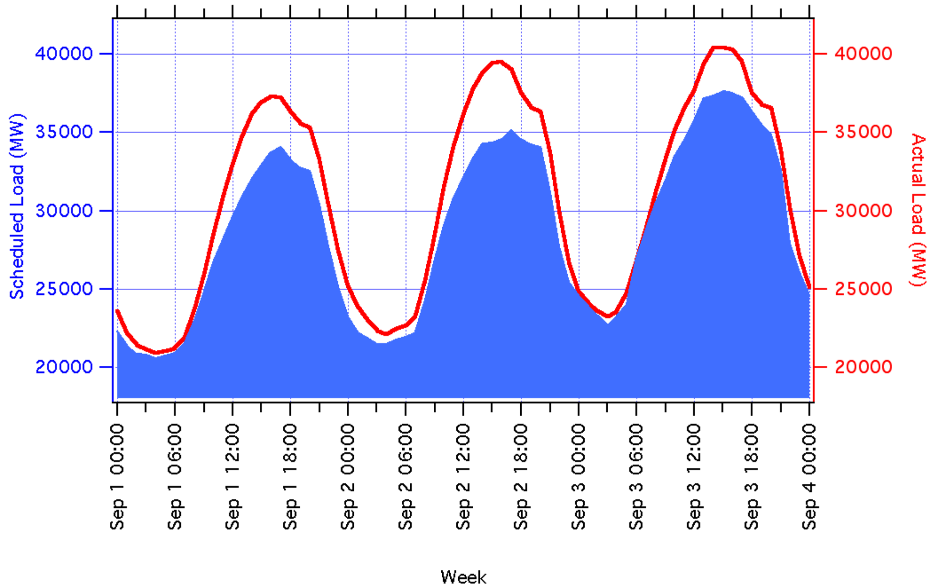

Figure 4.2. Scheduled and actual load over a three day sample period.

The difference between the scheduled load and the forecast load is the scheduling bias. It is typically negative (scheduled generation is less than forecast load) and, interestingly, reaches the largest negative values during peak summer hours when the power system is typically under the most stress. The scheduled load provided by the scheduling coordinators is often thousands of megawatts less than the forecast load provided by CalSO. Over the three year analysis period, the scheduled generation was as much as 5832 MW less than forecast load during peak hours. The average minima and maxima of the scheduling bias during peak hours are shown in Table 4.6 over the three year analysis period. The large negative bias of the hour ahead schedules provides an indication of the amount of generation assets available in the short term energy market. The data implies that the scheduling coordinators are comfortable with the depth of the generator stack; they can call up several thousand megawatts of generation whenever it might be needed. The scheduling bias was used as a proxy for estimating the depth of the generator stack. It was used for comparison purposes in determining the significance of renewable impacts on the system error.

The hour ahead schedules for each renewable generation resource were developed using a simple persistence model. This model provides a schedule of renewable output for the hour ahead market and is a conservative (worst-case) approach. Use of true forecasting models will reduce scheduling error and reduce the significance of renewable impacts from those calculated here. Figure 4.3 presents an example of actual output and scheduled output for a wind generator using the simple persistence model to calculate the schedule. The resource scheduling error was calculated as the difference between the resource's scheduled generation and its load following component of generation; with the hourly data used in this analysis, the hourly generation values were used directly as the value of the resource's load following component. The forecasting error including the scheduling error was then calculated by adding the resource scheduling error to the load forecasting error.

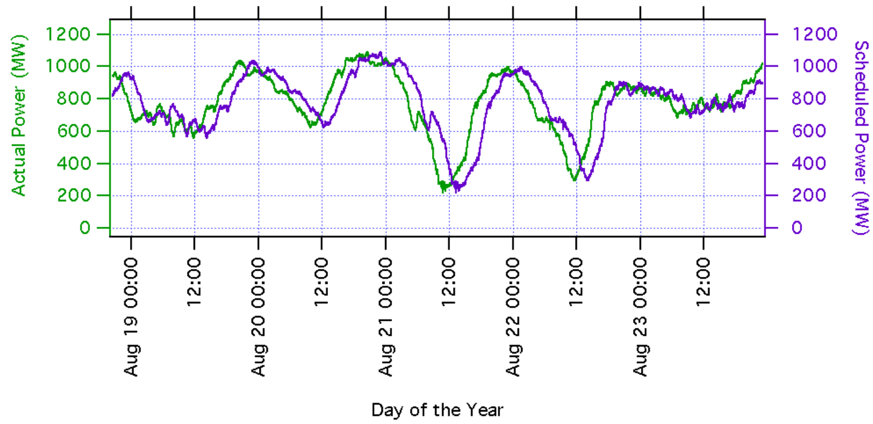

Figure 4.3. Actual and scheduled wind generation over a three day sample period. A simple persistence model was used to produce the schedule.

We compared the average minimum and maximum forecasting error during peak hours (noon to 6 p.m.) as a means of evaluating the significance of the renewable generator impacts. The results for the three analysis years are presented in Table 4.6. Negative values indicate that incremental energy purchases were required to compensate for under-generation or unexpected load. Positive values indicate over-generation or lower demand than expected, requiring generators in the short term energy market to decrement their output. The minimum forecasting error was changed by no more than two percentage points by any of the renewable resources with slight improvements in some cases. The impact on the maximum forecasting error was similarly small. This indicates that at current penetration levels, the scheduling error of the renewables do not have a significant effect on the total energy requirements from the short term market. The minimum scheduling bias reduced over the years but remained well over 200% greater than the load forecast error. This implies ample depth in the generator stack to handle incremental energy requirements.

Table 4.6. Results of multi-year analysis of forecast and scheduling errors during peak hours.

ERROR	2002				2003				2004			
	AVERAGE MINIMUM		AVERAGE MAXIMUM		AVERAGE MINIMUM		AVERAGE MAXIMUM		AVERAGE MINIMUM		AVERAGE MAXIMUM	
	MW	Compared to load forecast error alone (%)	MW	Compared to load forecast error alone (%)	MW	Compared to load forecast error alone (%)	MW	Compared to load forecast error alone (%)	MW	Compared to load forecast error alone (%)	MW	Compared to load forecast error alone (%)
Load forecast alone	-1945	100%	2112	100%	-1600	100%	2151	100%	-1439	100%	1529	100%
Load scheduling alone	-4747	244%	1302	62%	-4021	251%	2158	100%	-3700	257%	1776	116%
Scheduling bias	-5337	274%	1708	81%	-3336	208%	1534	71%	-3016	210%	1634	107%
Combined load forecast and renewable resource scheduling error												
Biomass	-1944	100%	2115	100%	-1603	100%	2157	100%	-1432	100%	1536	100%
Geothermal	-1947	100%	2112	100%	-1599	100%	2149	100%	-1442	100%	1529	100%
Solar	-1897	98%	2055	97%	-1631	102%	2153	100%	-1467	102%	1541	101%
Wind (Northern Cal)	-1946	100%	2148	102%	-1591	99%	2203	102%	-1419	99%	1554	102%
Wind (San Gorgonio)	-1930	99%	2142	101%	-1581	99%	2163	101%	-1443	100%	1545	101%
Wind (Tehachapi)	-1931	99%	2177	103%	-1569	98%	2181	101%	-1435	100%	1544	101%

4.4 Analysis of Ramping Capability

This analysis is presented as a complementary study to the forecast and scheduling error analysis above. It was originally developed by Brendan Kirby, ORNL and Michael Milligan, NREL.[11]

4.4.1 INTRODUCTION

In the load following time frame (from ten minutes up to a few hours), slow-start thermal generation that has already been committed (started so that sufficient resources are available to supply the expected load plus a reserve obligation) can be maneuvered to accommodate fluctuations in generation and system load. Combustion turbines or other fast-start units could be started in this time frame though that capability is not considered in this analysis (hourly availability data is not public). We assess the thermal generation load following capability that exists in the CaISO control area based on publicly available data. We then examine various renewable generation scenarios to determine if their load following requirements can be met with the capability supplied by the thermal units. This method is not as detailed as a full unit commitment and economic dispatch study, but can be useful for evaluating potential renewable impacts to load following costs.

4.4.2 SYSTEM RAMPING CAPABILITY

We estimated expected load following capability by examining the ramping capability of existing generators. Hourly load and generator data were obtained from Platts BaseCase, version 8.0.1. BaseCase provides hourly generation data for units that are subject to filing reports to the Environmental Protection Agency (EPA) for the Continuous Emissions Monitoring System (CEMS). This includes thermal generators, but hydro and nuclear units are not required to file and are therefore not represented in the database. Certain other generators are not required to file with CEMS, including some co-generation and some low-emission gas units. For the purposes of this study, the implication is that there is some existing generation in the control area that were not captured. Therefore, the hourly ramping capability was understated.

The 2002 data for the CaISO control area are presented in Table 4.7. In 2002, CaISO peak load was 42,352 MW. We obtained hourly data from 133 thermal generators representing a total capacity of 24,232 MW, which were included in the system ramping estimates. The 13,100 MW of hydro, 4,600 MW of nuclear, and 3,700 MW of other generation were not included in the ramping estimates. This discussion of the dataset limitation shows that our estimates of the CaISO control area's ability to ramp are understated, perhaps significantly. The results of our calculations and discussion below should therefore be interpreted as a minimum floor on the ramping capability that is available from thermal resources, and that capability can be complemented by other generation that we were unable to measure.

Table 4.7. Power requirements and generation mix of CaISO in 2002. Data from Platts BaseCase.

Load	
Peak load (MW)	42,352
Average load (MW)	26,573
Measured Thermal Generation	
Number of generators	133
Total capacity (MW)	24,232
Highest coincident output (MW)	17,541
Largest unit capacity (MW)	761
Average unit capacity (MW)	182
Average unit output (MW)	41
Additional Generation	
Hydro (MW)	13,100
Nuclear (MW)	4,600
Other (MW)	3,700

The first step in determining how much ramping capability is available and how much is needed is to determine the ramping capabilities of the individual generators. These capabilities are not publicly available, so we determined them by observing each generator's behavior. We analyzed a year of hourly generator output data to determine the maximum output, minimum non-zero operating output, and MW/min ramping capability for each generator. Generator maximum capability is simply the maximum hourly output the generator achieved during the year. Generator minimum capability and ramping capability are slightly harder to determine.

The minimum hourly output recorded in the data may be below the unit's actual sustained minimum operating capability. If the unit was turning on or off during the hour it would have spent part of the time at zero output, part of the time ramping on, and part operating stably. To better estimate the generator's minimum sustainable non-zero operating capability, we eliminate hours immediately after startup and immediately before shut down.

Each generator's ramping capability was determined by observing the maximum change in output between any 2 hours during the year. Upward and downward ramping were determined separately. As with the determination of the generator's minimum operating capability, hours immediately after startup and immediately before shutdown were excluded.

These estimates of generator capability are conservative. The generators may have greater capability that they simply did not have call to use during the year. Also, only hour-long ramps can be quantified. A 50 MW combustion turbine with a 20 MW minimum operating capability, for example, can be credited with a maximum 0.5 MW/min ramp rate, for example, regardless of the actual ramp rate capability. This is because the maximum change in output the unit can achieve is 30 MW and the evaluation interval is 60 minutes. The unit might be capable of ramping from 20 MW to 50 MW in under 10 minutes giving better than 3 MW/min ramp rate but the analysis methodology limits the calculated ramp rate to 1/6th that value. Conversely, this method does not capture other limitations such as temporary unit de-ratings or emissions limitations.

Knowing each generator's maximum and minimum operating capability and the up and down ramping capability allows us to determine the aggregate ramping capability available to the control area each hour of the year. System hourly MW/min ramping capability is the sum of the ramping capabilities of each generator that is on line that hour. Each generator's hourly ramping capability can be limited, for that hour, by the generator's current output and the maximum or minimum output capability. For example, a generator that is capable of 3 MW/min upward ramping would be limited to 0.2 MW/min if it had a maximum output capability of 200 MW and was operating at 188 MW during an hour (12 MW maximum ramp up / 60 minutes).

Table 4.8 summarizes the generator up and down ramping capabilities for the CaISO control area. The small size (182 MW) and the even smaller operating range of most units limits the calculated ramping capability for ramps lasting less than an hour. CaISO has a few large units that are also relatively fast. Again, these limitations combined with the unavailability of hydro data understates, in some cases significantly, the system ramping capability and correspondingly overstates the potential impact of nondispatchable renewables.

Table 4.8. Thermal generator ramping capabilities, CaISO in 2002.

Measured thermal generation	Ramping capability (MW/min)
Fastest unit MW/min ramp capacity (up/down)	8.6 / -7.8
Average unit MW/min ramp capacity (up/down)	1.6 / -1.6
Total capacity (up/down)	215 / -214
Total simultaneous capacity (up/down)	168 / -175
Maximum used capability (up/down)	42 / -66

The ramping capability available to the control area is the sum of the individual generators' ramping capabilities. This aggregate capability varies from hour to hour as different generators come on and off and as their operating levels vary. Having determined the maximum and minimum output along with the ramping capabilities of each generator we were able to reexamine the year of load data and determine, for each hour, what the control area ramping requirements were and what excess ramping capability was available from the thermal generation. We only consider generation ramping capability that is in the same direction as the current load requirement. That is, up-bound ramping capability is evaluated when the load is ramping up and down-bound ramping capability is evaluated when the load is ramping down.

The last three lines of Table 4.8 present the total control area thermal ramping capabilities. As expected, the total capability of all the units exceeds the maximum capability that is ever actually available. There are two primary reasons for this. First, all the units are never on line at the same time. Second, some of the thermal units are typically operating near their full output so they have limited capability to ramp up. Significantly, the full capability was never used during the year.

4.4.3 LOAD RAMPING REQUIREMENTS

Thermal ramping capabilities typically exceeded the control area load ramping requirements. Figure 4.4 presents histograms of both the generation capabilities and the load requirements. The CaISO control area tends to operate with generators partially loaded for many hours of the year. Generators are poised to move up or down and the generation ramping capabilities histogram is fairly symmetric.

The histograms presented in Figure 4.4 do not show simultaneous requirements and capabilities. Figure 4.5 presents simultaneous load ramping requirements and thermal generation ramping capability as a ratio. Thermal ramping capability exceeds load requirements, in both the up and down directions, for all but 100 hours. For most hours the thermal ramping capability far exceeds the load ramping requirements. The extremely high ratios of capability to requirements on the left side of the graph result from times when the load is not ramping much and are not overly significant. The excess capability represented for many hours in the middle of the graph, when the load is ramping moderately, are more important. The control area never fell short of ramping capability; significant hydro and other ramping resources are available to the control area but are not captured in our data.

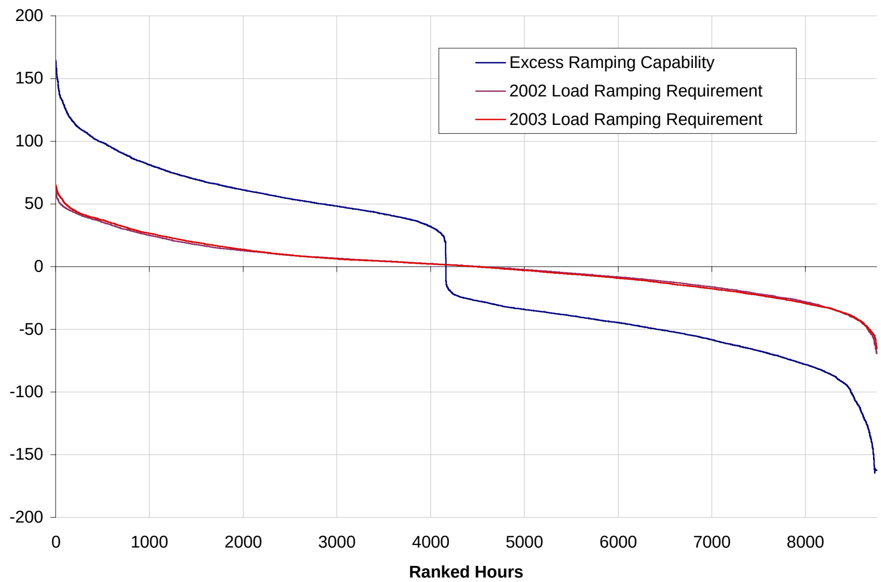

Figure 4.4. Thermal ramping capability and load ramping requirement. The 2002 and 2003 load ramping requirement traces (purple and red) are nearly overlapped. Thermal ramping capabilities typically exceed load ramping requirements.

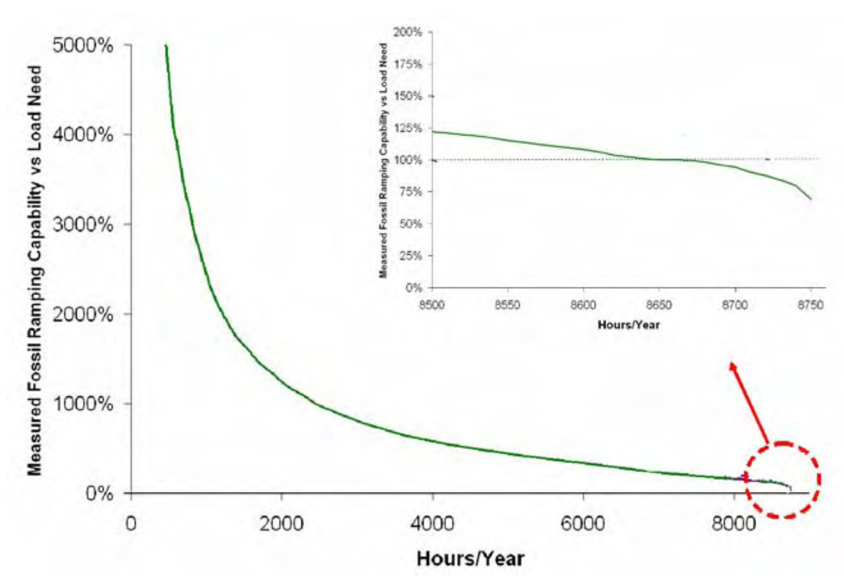

Figure 4.5. The ratio of simultaneous load ramping requirements and thermal generation ramping capability, 2002. Thermal ramping capability exceeds load ramp requirements more than 97% of the time.

4.4.4 RENEWABLE GENERATOR RAMPING REQUIREMENTS

Some renewable generators have a time varying output, which must be countered by conventional units. The ramping requirement of these renewables represents the rate of change that is required from conventional units to compensate for the renewables' variations in the load following time frame. Ramping requirements were calculated using the CaISO one year 2002 dataset. They were also calculated using the higher quality hourly datasets from the IOUs; however, hourly data, as discussed above, understates ramping requirements.

Wind and solar generation have the largest ramping requirements and were the focus in this effort. The total wind ramping requirement for 2002 is shown in Figure 4.6. During this year, the maximum wind generation was approximately 1200 MW. The peak ramp-up requirement for wind occurred during the month of May (Figure 4.7), while the peak ramp-down requirement was during the month of February (Figure 4.8). Diurnal ramping requirements for wind generation during the summer months were typically less than 7 MW/minute, as shown in Figure 4.9 and Figure 4.10.

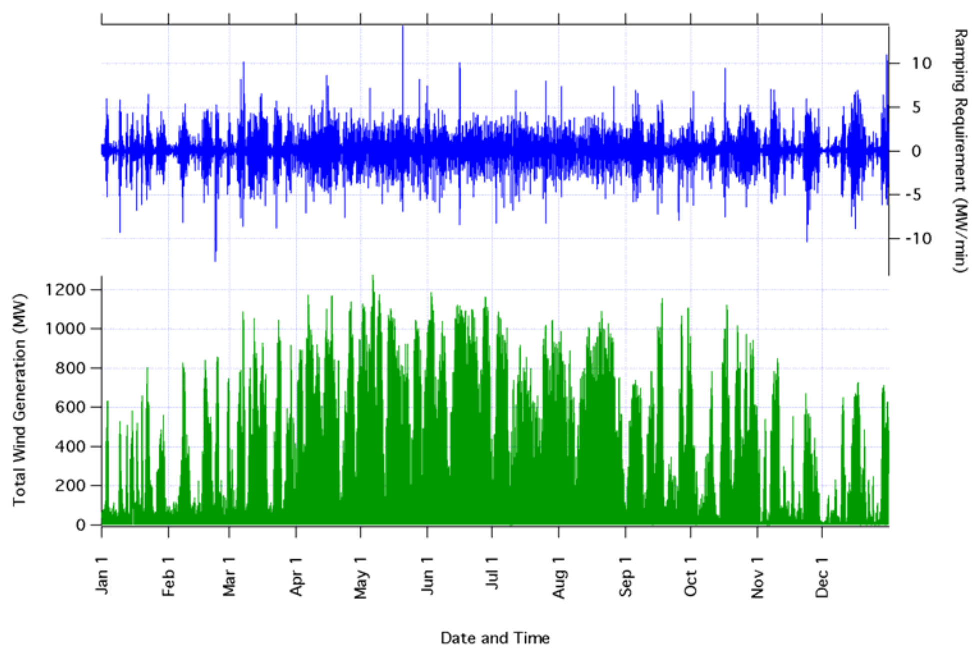

Figure 4.6. The total wind ramping requirement in California, 2002, calculated from ten minute averages.

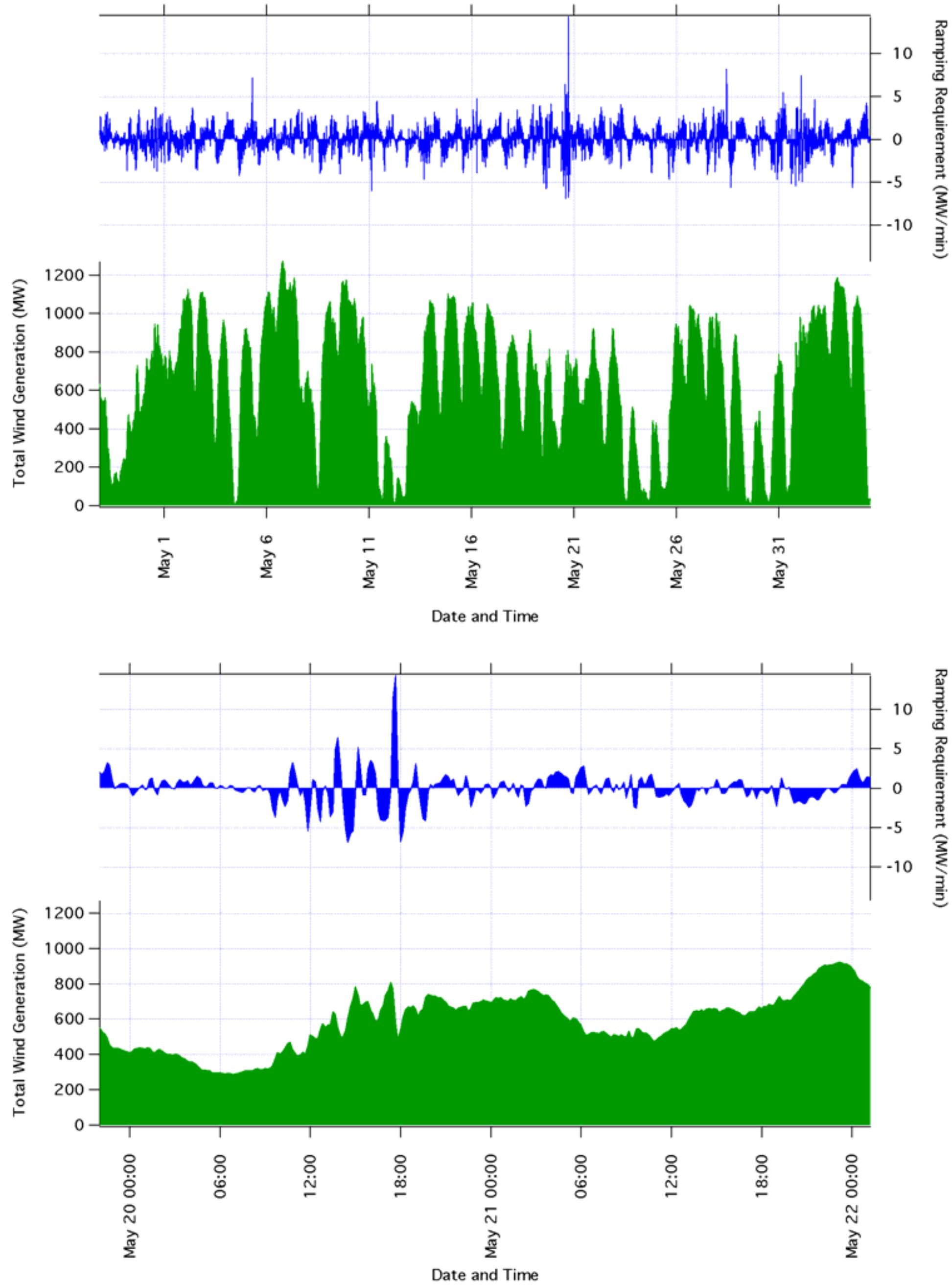

Figure 4.7. The total wind ramping requirement in May 2002, showing large ramp-up requirements.

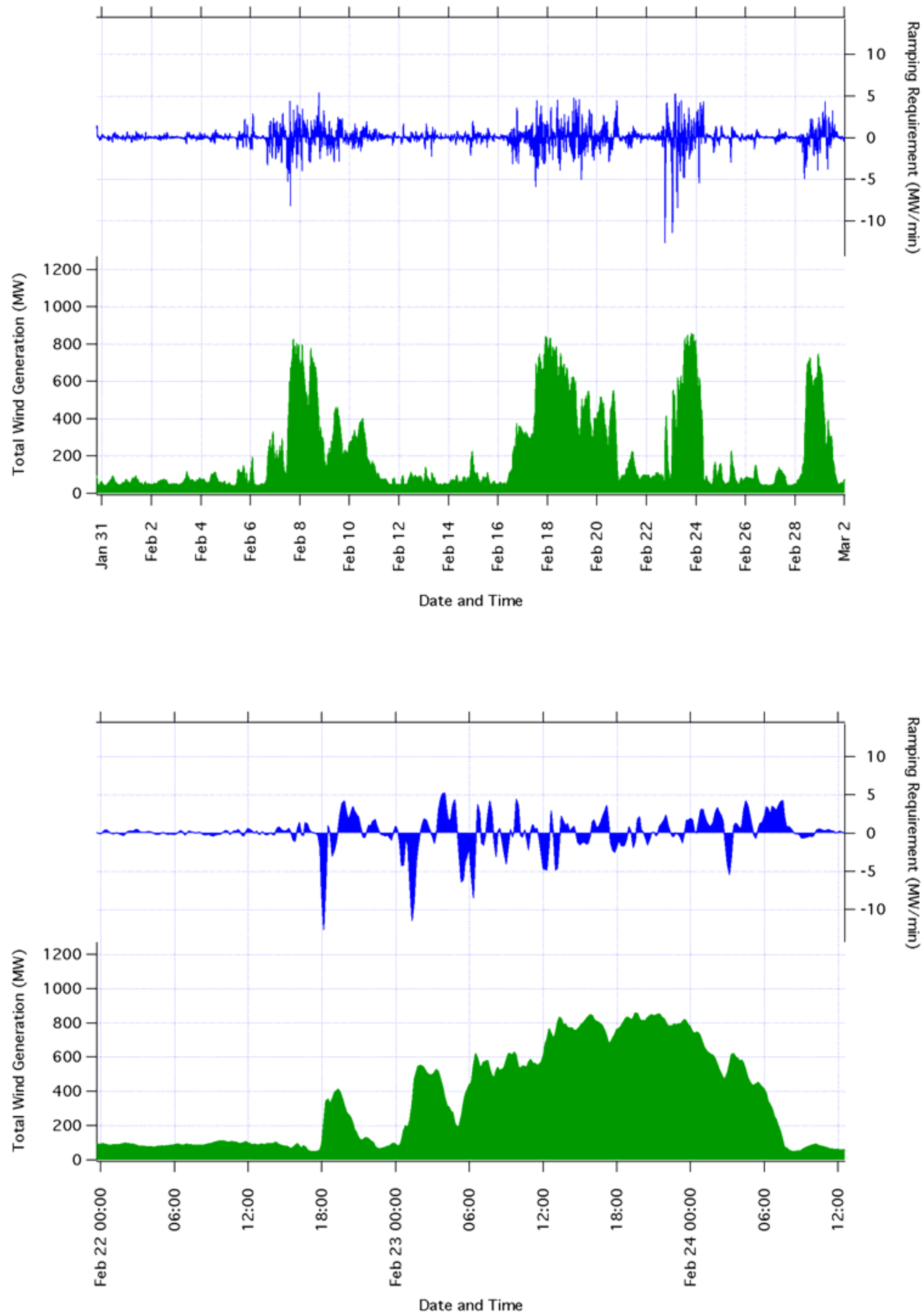

Figure 4.8. The total wind ramping requirement in February 2002, showing large ramp-down requirements.

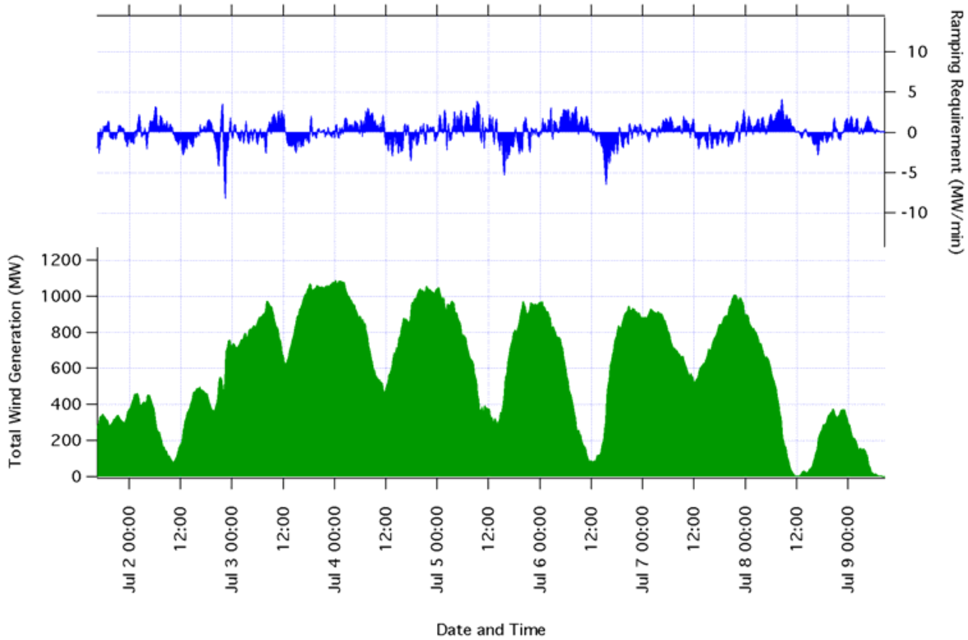

Figure 4.9. Typical total wind ramping requirements, shown in July 2002.

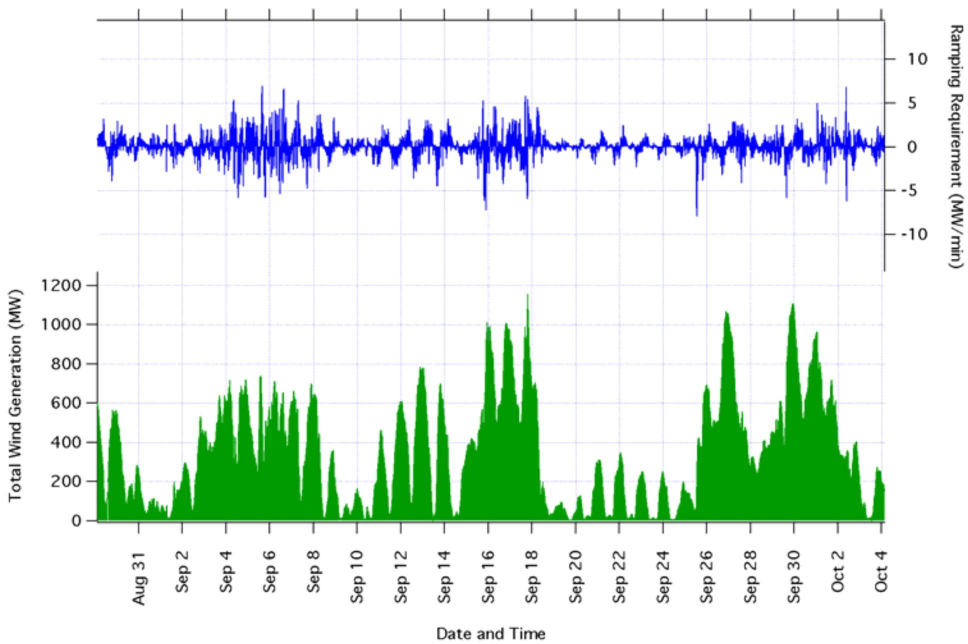

Figure 4.10. Typical total wind ramping requirements, shown in September 2002.

The ramping requirements for solar generation were evaluated using 10 minute averages for the 2002 analysis year (Figure 4.11). The maximum solar generation was approximately 350 MW during this year. Solar generation has a diurnal pattern which requires ramping in the morning and evening hours (Figure 4.12 through Figure 4.14).

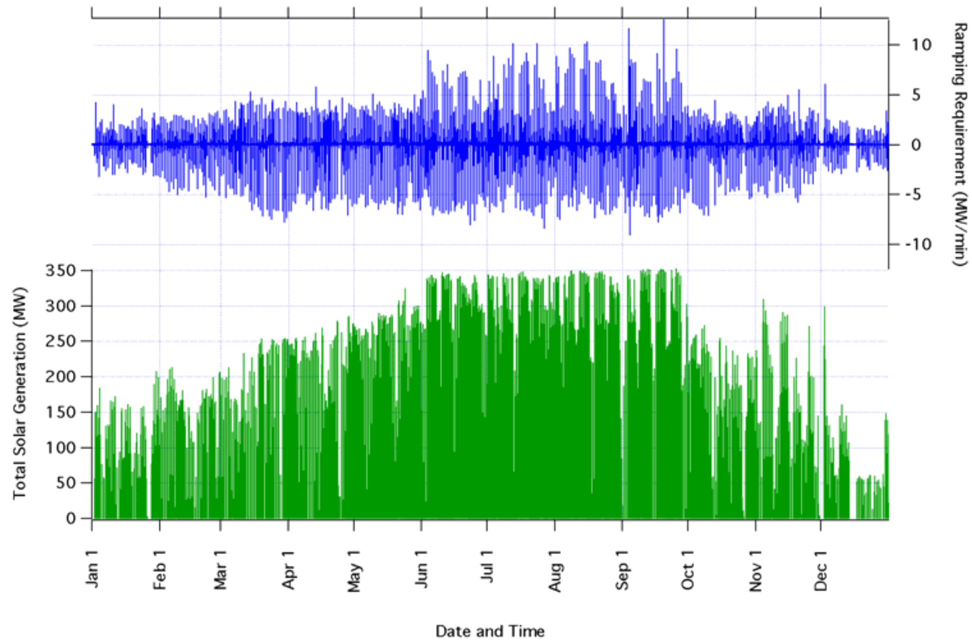

Figure 4.11. Solar ramping requirement in California, 2002, calculated from ten minute averages.

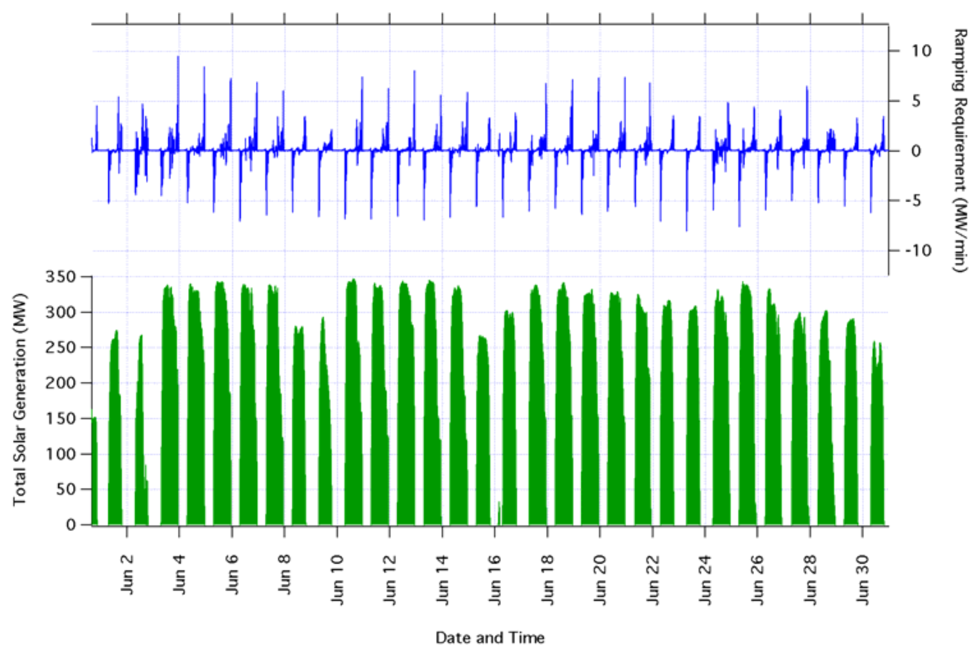

Figure 4.12. Solar ramping requirements, June 2002.

Figure 4.13. Solar ramping requirements in July 2002, showing large ramp-down requirements.

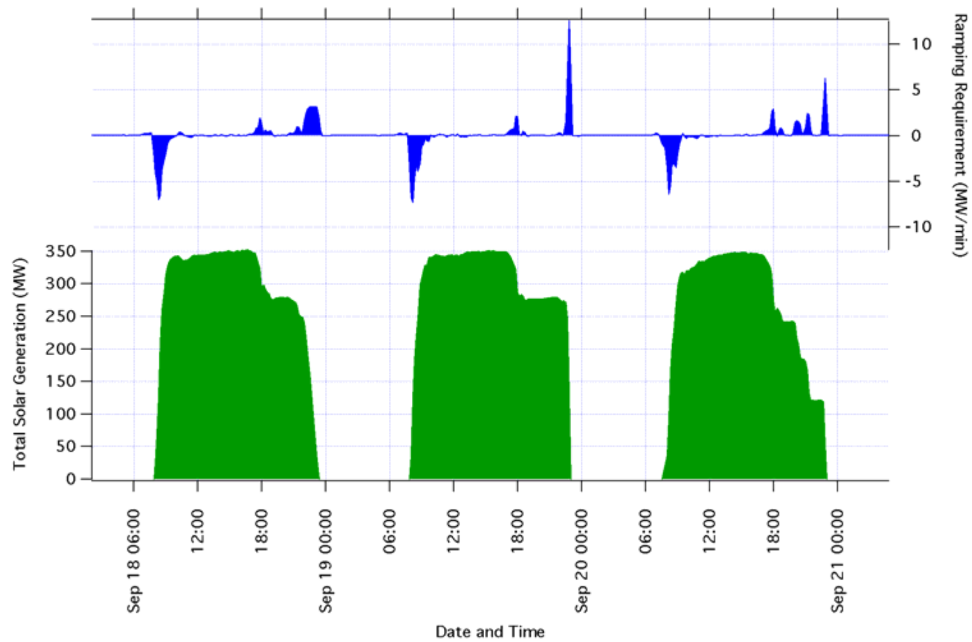

Figure 4.14. Solar ramping requirements in September 2002, showing a large ramp-up requirement.

A comparison was made to determine the impact of calculating ramping requirements based on hourly data rather than 10 minute data. The results showed that calculating ramping requirements using the hourly data resulted in similar trends, as shown in

Figure 4.15. The hourly ramping requirements for wind and solar during 2002 were negligible compared to the load ramping needs (Figure 4.16). The ramping requirements were also found to be consistent from year-to-year as shown in Figure 4.17 and Figure 4.18.

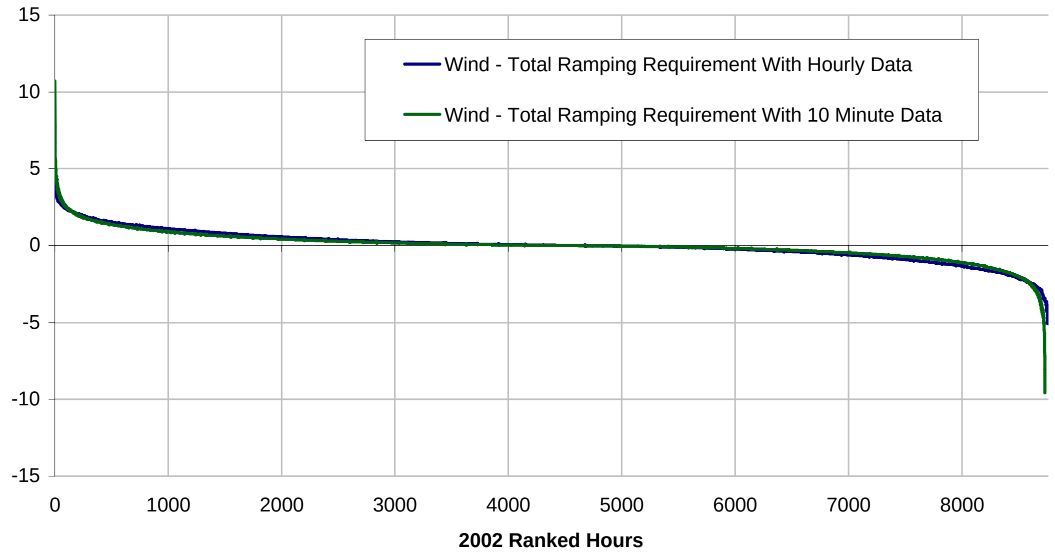

Figure 4.15. Comparison of wind ramping requirements calculated with 10 minute and hourly data.

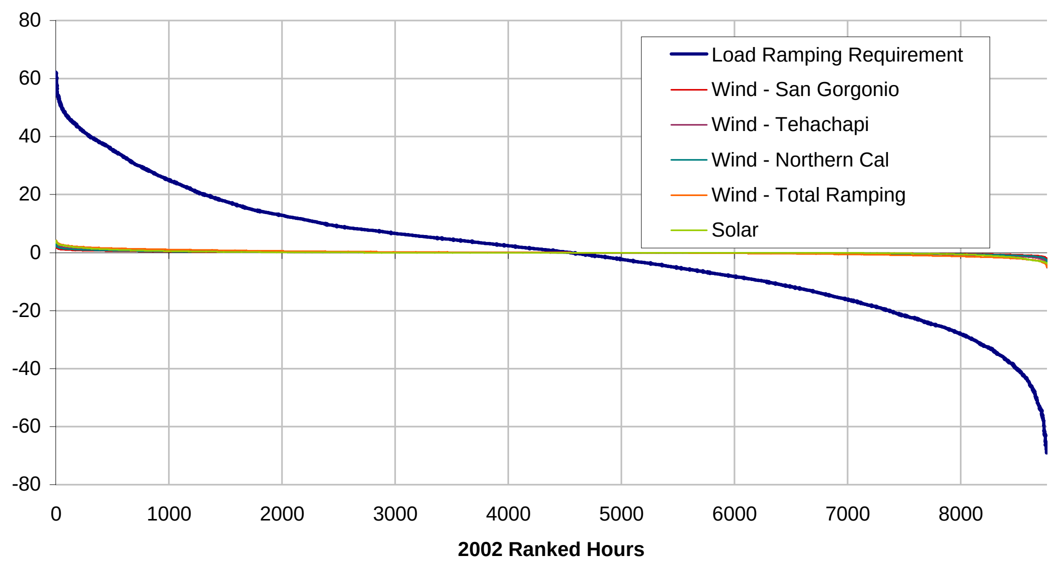

Figure 4.16. Ramping requirements for wind and solar aggregates based on hourly data; the requirements are small compared to the load ramping requirement.

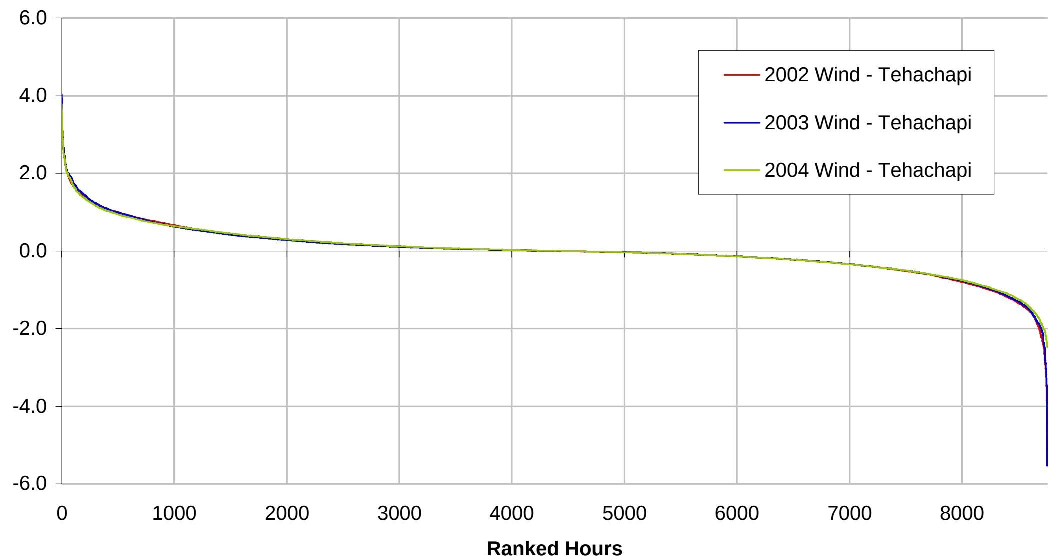

Figure 4.17. The ramping requirement of wind in Tehachapi based on hourly data over three years.

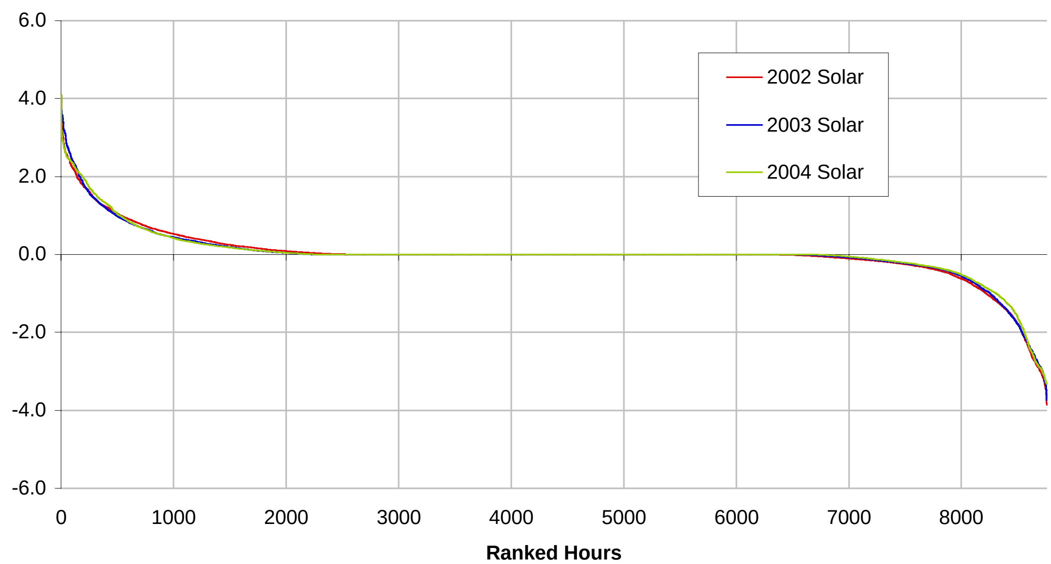

Figure 4.18. The ramping requirement of solar based on hourly data over three years.

4.4.5 DISCUSSION OF THE RAMPING CAPABILITY ANALYSIS RESULTS

It is possible to calculate a lower bound to the ramping capability within a given control area using public databases. In our experience some significant capabilities could not be estimated and more ramping capability exists than we were able to measure.

It appears that there is a very large amount of ramping capability in the CaISO control area during most hours of the 2002 analysis year we studied. This ramping capability is a natural result of the resource mix that has developed. Because each increase or decrease of renewable generation does not need to be matched one-for-one by another generator, the ability to absorb moderate or even large quantities of wind, solar, and other renewables appears significant for most of the year.

The CaISO control area appears to have significant ramping resources available from thermal generation that is partially loaded and physically able to respond. CaISO, like most ISOs, operates energy markets that clear several times an hour, providing access to the ramping capabilities of the generators active in the energy markets. Control areas that do not have access to fluid intra-hour markets still have the physical capabilities of the generators but may not have *access* to that capability simply based on the hourly market structure. This lack of access denies the generators the ability to position themselves (ramp) to sell as much energy as customers want, forces the control area operator to use additional regulating resources instead, and forces consumers to pay for the inefficiency.

There may be significant opportunities for neighboring control areas to assist each other in the load following time frame as well. This is partly a natural consequence of the ability of larger control areas to better manage variability, whether caused by load, wind, or a combination with other resources. It is also a consequence of additional capability being inherently available from a larger pool of generators.

Assessing the ramping capability of a control area with public data presents some challenges. Because some data are unreported, and because of the shortcomings of our method, it is not possible to obtain an accurate measure. However, having said that, we think that this type of analysis can be useful in several ways. The estimates provided by this approach provide a lower bound on the load following capability in a control area. The approach is transparent, which makes it possible to more easily understand how the more complex methods embodied in production simulation models work. The approach could easily be extended to include data from non-CEMS-reporting resources. For entities that have access to such data, a more detailed analysis would be possible, and would provide a better estimate of the load following capability of the control area.

5 DATA

5.1 Requirements

As detailed in the previous sections, the integration cost analysis requires a broad array of input data. The minimum data requirements of the analysis are listed in the following table.

Table 5.1. Minimum input data requirements for integration cost analysis.

Data item	Units	Sampling rate
Generation of each resource being studied	MW	One minute, hourly
Load	MW	One minute, hourly
Interchange	MW	Hourly
Hydro generation	MW	Hourly (preferred)
Generation supply data of all other generators in the study area, including capacity and outage rates	MW; outage rates are expressed as fractions	Hourly
Regulation up, procured	MW-hr	Hourly
Regulation down, procured	MW-hr	Hourly
Regulation up, self-provided	MW-hr	Hourly
Regulation down, self-provided	MW-hr	Hourly
Price of regulation up	$/MW-hr	Hourly
Price of regulation down	$/MW-hr	Hourly
Hour ahead load forecast	MW	Hourly
Hour ahead load schedule	MW	Hourly

While not essential to the completion of the analysis, additional generation and system operation data can be beneficial for higher fidelity modeling and data verification.

Because the accuracy of the analysis results and the input data are tightly coupled, high quality input data is necessary. As discussed in the next section, a variety of data sources were used to develop an input dataset of sufficient quality for the analysis.

5.2 Datasets

Numerous data sources were used to compile the input dataset for the multi-year analysis, including public and proprietary datasets from CaISO, California's IOUs, and a commercial database. At the onset of the study, it was assumed that CaISO would provide all the data necessary for the analysis. As the study progressed, additional sources were sought when particular data was unavailable from CaISO or to address data quality issues. The various datasets used are listed and described below in the chronological order in which they were incorporated into the study. Issues encountered in the datasets and the methods used to address them are discussed below in Section 5.3.

5.2.1 CAISO OASIS HOURLY DATA

CaISO provides a web accessible, publicly available database known as the Open Access Same-Time Information System or OASIS. OASIS contains current and archived market data for energy and transmission in California including actual, scheduled, and forecasted load values and actual regulation purchase amounts and prices. It can be accessed at http://oasis.caiso.com/.

OASIS data was used in both the previous one year analysis and the multi-year analysis detailed in this report. A number of scripts were developed to automate the retrieval and collation of the data. The OASIS data used in the multi-year analysis is listed and described in the following table.

Table 5.2. OASIS data used in the multi-year analysis.

Data item	Units	Sampling rate
Regulation up, procured, pre-rational buyer	MW-hr	Hourly
Regulation down, procured, pre-rational buyer	MW-hr	Hourly
Regulation up, self-provided	MW-hr	Hourly
Regulation down, self-provided	MW-hr	Hourly
Price of regulation up, procured, pre-rational buyer	$/MW-hr	Hourly
Price of regulation down, procured, pre-rational buyer	$/MW-hr	Hourly

5.2.2 CAISO ONE-YEAR 2002 DATASET

The CaISO one-year 2002 dataset, often referred to simply as the one-year dataset, consists of system operation and power generation data for 2002 sampled at a one minute interval. It was used previously in the Phase I one-year analysis and is detailed in the Phase I report. Unlike the OASIS data, the one-year dataset is not publicly accessible and was released for the integration cost study through nondisclosure agreements. The dataset includes the following:

Table 5.3. CaISO one year 2002 dataset.

Data item	Annual peak, MW (where appropriate)	Notes
Load, total	42,388	
Generation, total		
ACE		
Interchange, actual		
Interchange, scheduled		
Frequency, actual		
Frequency, scheduled		
Regulation, total		
Deviation from preferred operating point		
Biomass generation aggregate	413	
Geothermal generation aggregate	155	
Solar generation aggregate	352	Includes gas assist generators.
Wind generation aggregate, Altamont	437	
Wind generation aggregate, Pacheco		
Wind generation aggregate, San Gorgonio	287	
Wind generation aggregate, Solano		
Wind generation aggregate, Tehachapi	578	
Wind generation aggregate, total		Calculated sum of above regional aggregates
Wind generation aggregate, total 2		Slightly different than the wind total above; data was recorded from a different source.
Conventional generation aggregates		A variety of conventional generation aggregates including gas-fired steam units, combined cycle units, untis on automated dispatch, and units on AGC.

Although non-aggregated data was initially requested for the study, CaISO was able to provide only aggregated generation data because of confidentiality concerns, even with nondisclosure agreements. Data was aggregated by generation subsets based on renewable resource type and, in the case of wind, by region. Although each aggregate does not comprehensively include all of the selected generators of interest within CaISO's control area, CaISO attempted to include sufficient capacity within each aggregate so that it could be representative.

The data was extracted by CaISO from their Plant Information (PI) system, a vast, internal database of system operation and power generation data for California. The PI system contains over 180,000 data fields, commonly referred to as PI tags, and finding the appropriate PI tags for the desired data was a nontrivial task. Extracting the data was also nontrivial, requiring both extensive computer time and manual intervention to ensure complete extractions.

Once the raw data was extracted and aggregated by CaISO, it was collated and manually reviewed for errors by the Integration Cost Study analysis team. Data spikes and dropouts, detailed in Section 5.3.6, were removed. During the data review and error correction process, it was found that data aggregation introduces significant difficulties in detecting data quality issues, as discussed further below.

5.2.3 BASECASE DATA

Platts (formerly Resource Data International or RDI) BaseCase is a commercial database of power market and systems data. BaseCase was used to identify non-renewable generators and their capacities and the forced outage and maintenance rates of non-intermittent generators. This data was used in both the one-year analysis and the multi-year analysis.

5.2.4 CAISO MULTI-YEAR DATASET

A new dataset was provided by CaISO for the multi-year analysis. Like the one-year dataset, the multi-year dataset contains system operation and power generation data sampled at one minute intervals from CaISO's PI system. It covers January 1, 2002 to mid-September 2004 and was released for this study through a confidentiality agreement. The specific items in the dataset are listed in the table below.

Table 5.4. CalSO multi-year dataset.

Data item	Annual peak, MW (where appropriate) 2002	2003	2004	Notes
ACE				
Load	42388	42671	45582	
Frequency deviation				
Interchange, actual				
Interchange, scheduled				
Interchange deviation				
Calculated ACE				
Difference between actual and calculated ACE				
Biomass generation aggregate	462	460	473	Data gap from Feb 12, 2002 to Sep 17, 2002.
Solar generation aggregate	352	352	350	Includes gas assist generators. Data gap from Feb 13, 2002 to Sep 18, 2002.
Wind generation aggregate, Altamont	445	545	615	The peak generation values shown here are totals of the Altamont, Pacheco, and Solano aggregates.
Trustworthy aggregated capacity, Altamont				
Wind generation aggregate, Tehachapi	578	579	572	
Trustworthy aggregated capacity, Tehachapi				
Wind generation aggregate, San Gorgonio	287	381	513	
Trustworthy aggregated capacity, San Gorgonio				
Wind generation aggregate, Pacheco				See Altamont, above.
Wind generation aggregate, Solano				See Altamont, above.
Trustworthy aggregated capacity, Solano				
Wind generation aggregate, calculated				Calculated sum of above regional aggregates
Wind generation aggregate, EMS total				Different than the calculated wind total above; data was recorded from the EMS system.
Geothermal generation aggregate, SCE territory	144	338	356	
Trustworthy aggregated capacity, geothermal, SCE territory				
Geothermal generation aggregate, QF total				
Geothermal generation aggregate, Geysers				
Trustworthy aggregated capacity, geothermal, Geysers				
Generation of a combined cycle gas unit				
Trustworthy aggregated capacity, of above combined cycle gas unit				

While similar in construct, there are several key differences between the one-year and multi-year datasets. The composition of the generation aggregates is not identical between the two datasets. The aggregates were expanded in the multi-year dataset to include more generators. This was intended to make the aggregates more representative of their generation subsets. Because many of the aggregates from the one-year dataset already included most of the generators that they represented, the capacities of the aggregates in the multi-year dataset are not all significantly higher. However, the inclusion of additional PI tags in some cases introduced significant new data issues with elevated data floors and ramps, as detailed in Section 5.3.7.

To address the data issues encountered in the one-year dataset, CaISO included a *trustworthy aggregated capacity* for several of the generation data streams in the multi-year dataset. The PI system is able to detect some types of data errors such as telemetry errors and records a data quality tag along with incoming data values. There are also errors that the PI system does not automatically detect such as metering errors, some errors in data received by the PI system that is already aggregated, and complete data dropouts. CaISO developed an error detection method that combines the PI quality tags and an algorithm which uses timestamps of recorded incoming data. For each aggregated generation data value, the method provides the summed capacity of the generators in the aggregate that are found to have good data; the summed capacity of the generators reporting good data is referred to as the *trustworthy aggregated capacity*. Unfortunately, the values of the trustworthy aggregated capacity did not correlate well with manually detected errors in the generation data and another method for addressing the issues in the multi-year dataset was deemed necessary.

Two other methods were considered and pursued simultaneously, as described in Section 5.3.8. Ultimately, generation data from the IOUs was used to resolve the data quality issues. Hourly generation data provided by PG&E and SCE was of sufficiently high quality that it was used directly in the capacity credit and load following analyses in lieu of data from the CaISO multi-year dataset. However, neither PG&E nor SCE record data at a rate fast enough for the regulation analysis. As described in Section 5.3.8, the IOU data was used as a basis to identify errors in the raw CaISO multi-year dataset. The processed and corrected multi-year dataset was then used in the regulation analysis.

5.2.4.1 Discussion of Renewable Generation Data

In the plots of power generation data excerpts below, the axis values have been intentionally left off to preserve data confidentiality.

Figure 5.1 and Figure 5.2 show the output of the biomass aggregate in both long and short timescales. Generation almost drops to zero in fall of 2002 and in spring of 2003 and 2004. There are no clear diurnal generation patterns; output is sometimes fairly constant, sometimes moves in blocks as if on a scheduled dispatch, and sometimes moves without obvious reason. The biomass data spans January 1, 2002 to late September 2004, except for a period in 2002 from mid-February to mid-September in which data is missing.

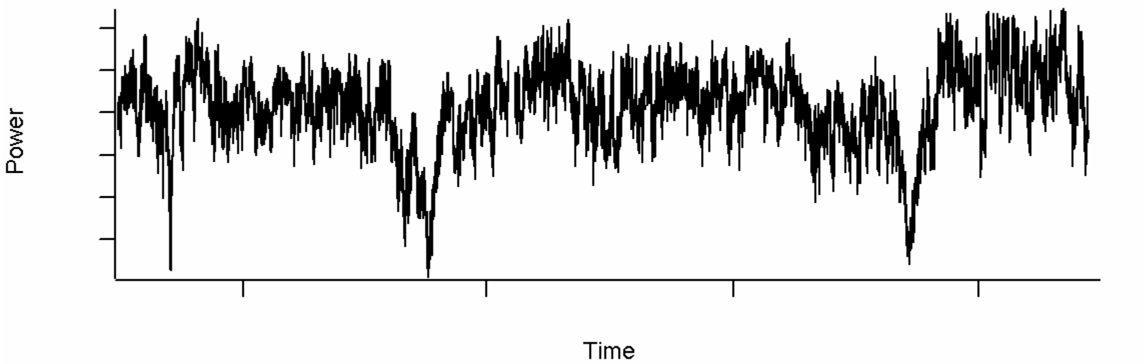

Figure 5.1. Generation of the biomass aggregate in the CalSO multi-year dataset. Two years from fall of 2002 to fall of 2004.

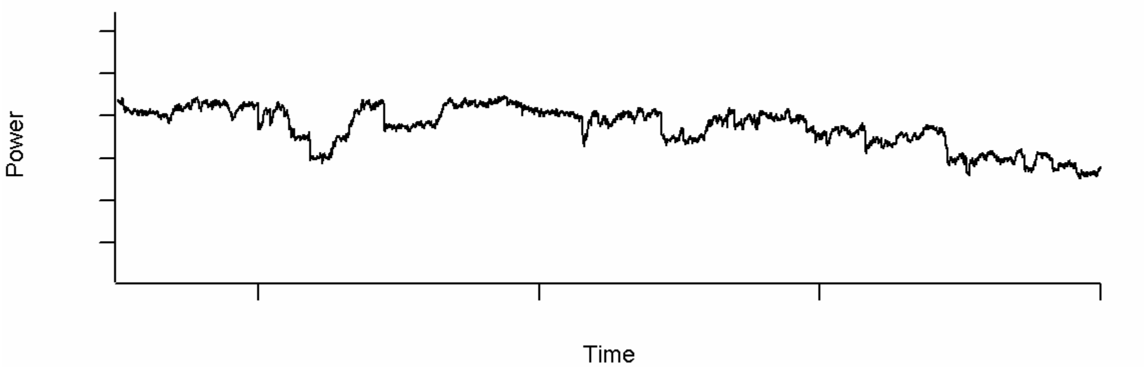

Figure 5.2. Generation of the biomass aggregate in the CalSO multi-year dataset. One week from winter of 2004.

CaISO provided two aggregates of geothermal generation in the multi-year dataset, one for the Geysers region and one for SCE territory. The Geysers data did not match well with any of the IOU generation data and could not be reviewed for errors as described in Section 5.3.8. Consequently, CaISO's geothermal data for SCE territory was used in the analysis, but not the CaISO Geysers data (recall that CaISO generation data was used only in the regulation analysis and that IOU generation data, not CaISO generation data, was used in the capacity credit and load following analyses).

Figure 5.3 and Figure 5.4 show the output of the geothermal aggregate for SCE from the CaISO multi-year dataset. Two noteworthy occurrences appear in the data. As seen in Figure 5.3, prior to May 2002, the data shows block scheduling with distinct morning and evening ramps on weekdays. Afterwards, the aggregate exhibits relatively constant output except for occasional drops until spring 2003 when, as shown in Figure 5.4, the power output more than doubles.

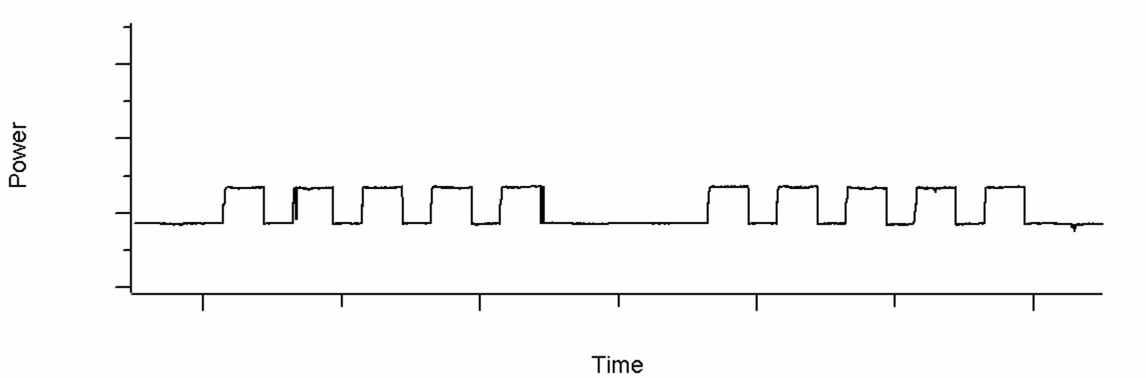

Figure 5.3. Generation of the SCE territory geothermal aggregate in the CalSO multi-year dataset. Two weeks in winter of 2002.

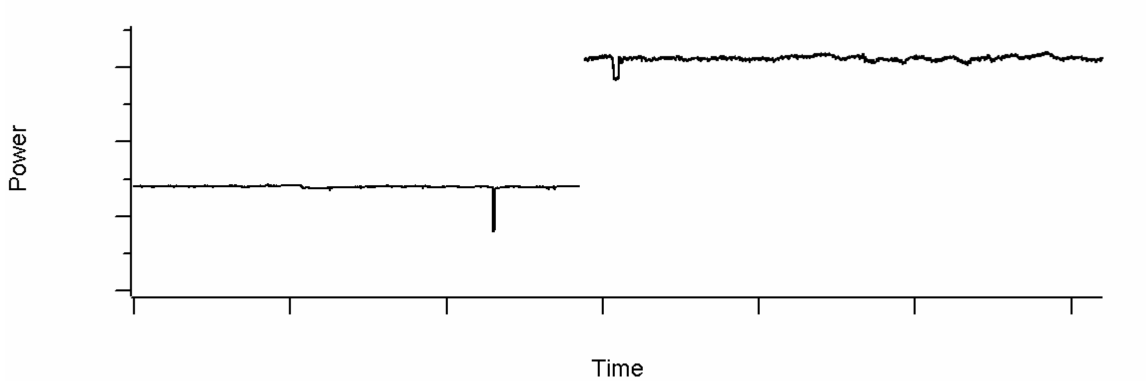

Figure 5.4. Generation of the SCE territory geothermal aggregate in the CalSO multi-year dataset. One month in spring of 2003.

Figure 5.5 and Figure 5.6 show the output of the solar aggregate over one year and one month. Seasonal and diurnal trends are strongly evident. California's large solar plants have gas generators to augment the power produced by their solar concentrators/collectors; the data includes the power output from these gas generators. The data spans January 1, 2002 to late September 2004, except for a period in 2002 from mid-February to mid-September in which data is missing.

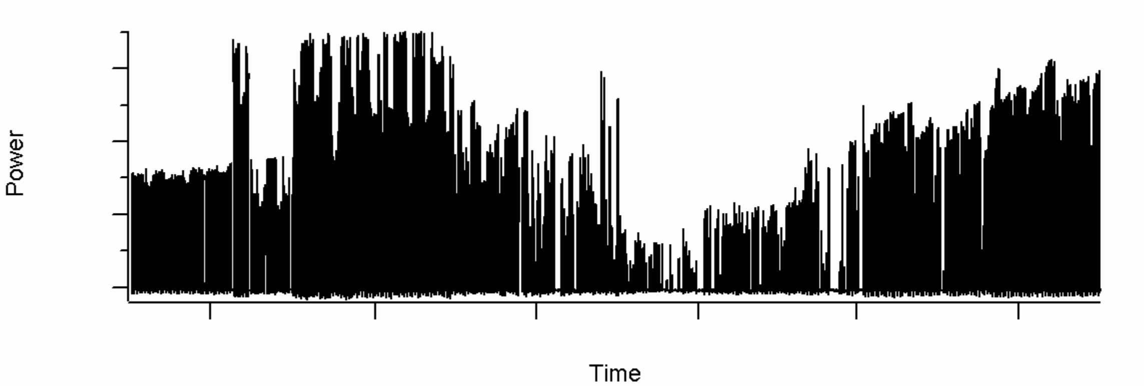

Figure 5.5. Generation of the solar aggregate in the CalSO multi-year dataset. One year from summer of 2003 to summer of 2004.

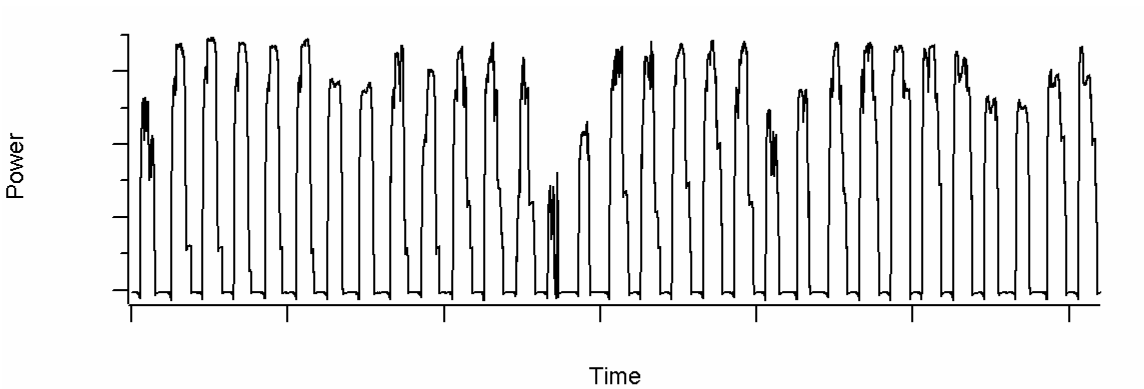

Figure 5.6. Generation of the solar aggregate in the CalSO multi-year dataset. One month in summer of 2004.

CalSO provided regional wind generation aggregates for the Altamont, Pacheco, San Gorgonio, Solano, and Tehachapi areas. PG&E provided a single wind aggregate that contained plants from Altamont, Pacheco, and Solano. To perform a data review against corresponding PG&E data, the CalSO Altamont, Pacheco, and Solano generation data were combined to a form an aggregate that encompasses all three of Northern California's largest wind resource areas. Figure 5.7 and Figure 5.8 show the power output of this aggregate. Figure 5.9 through Figure 5.12 show the output of the San Gorgonio and Tehachapi wind aggregates. Seasonal and diurnal trends are strongly evident in all the wind generation data. The Northern California wind aggregate, as shown in Section 5.3.7, exhibited the most problems with elevated data floors.

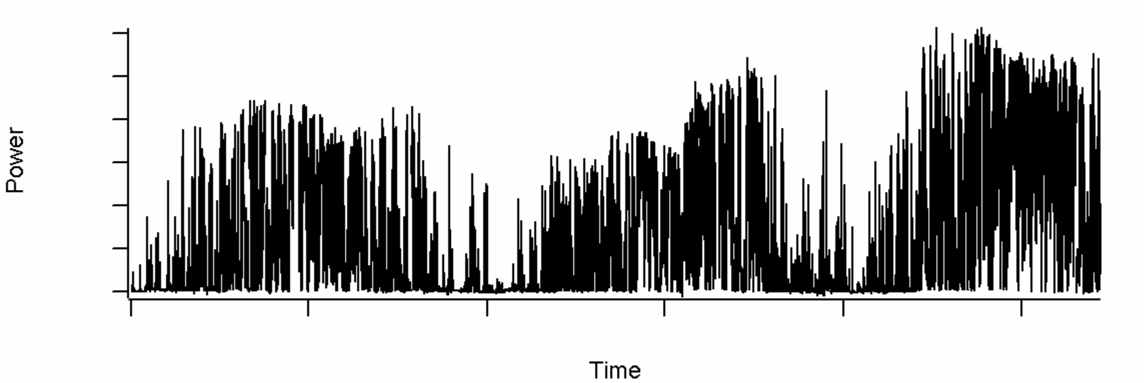

Figure 5.7. Generation of the Northern California (Altamont, Pacheco, Solano) wind aggregate in the CalSO multi-year dataset. January 2002 to September 2004.

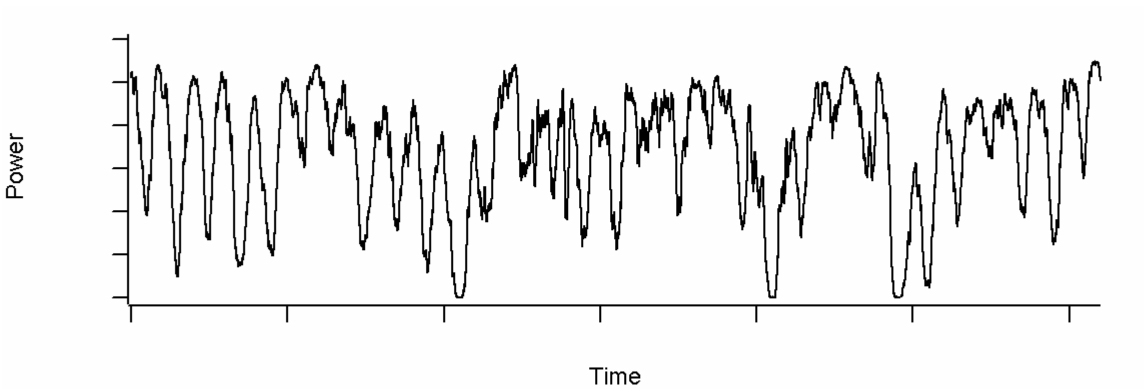

Figure 5.8. Generation of the Northern California (Altamont, Pacheco, Solano) wind aggregate in the CalSO multi-year dataset. One month in summer 2004.

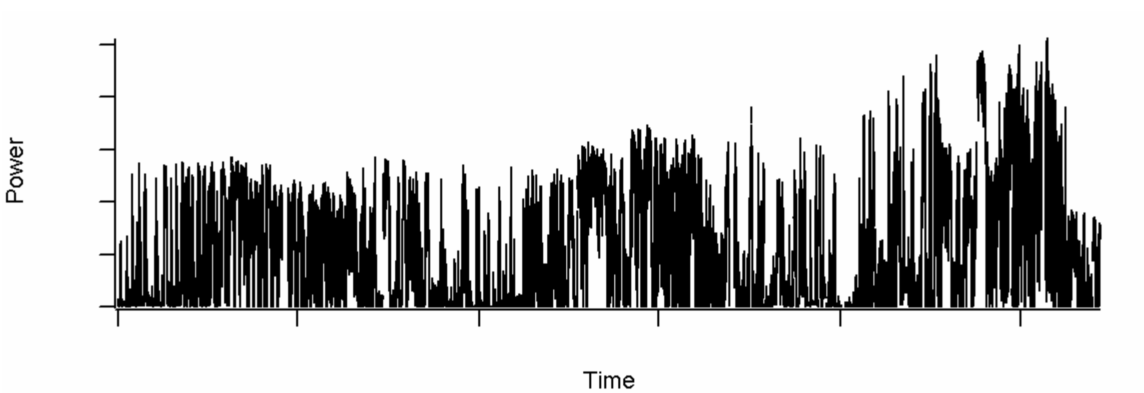

Figure 5.9. Generation of the San Gorgonio wind aggregate in the CalSO multi-year dataset. January 2002 to September 2004.

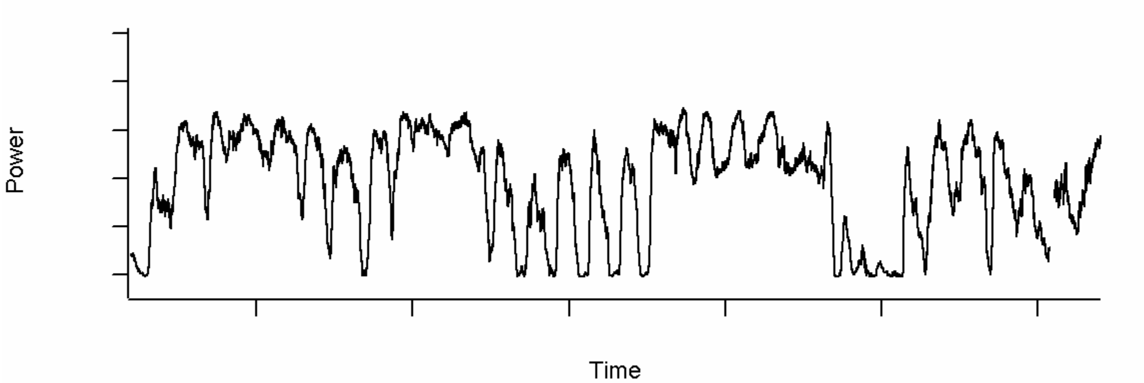

Figure 5.10. Generation of the San Gorgonio wind aggregate in the CalSO multi-year dataset. One month in summer of 2003.

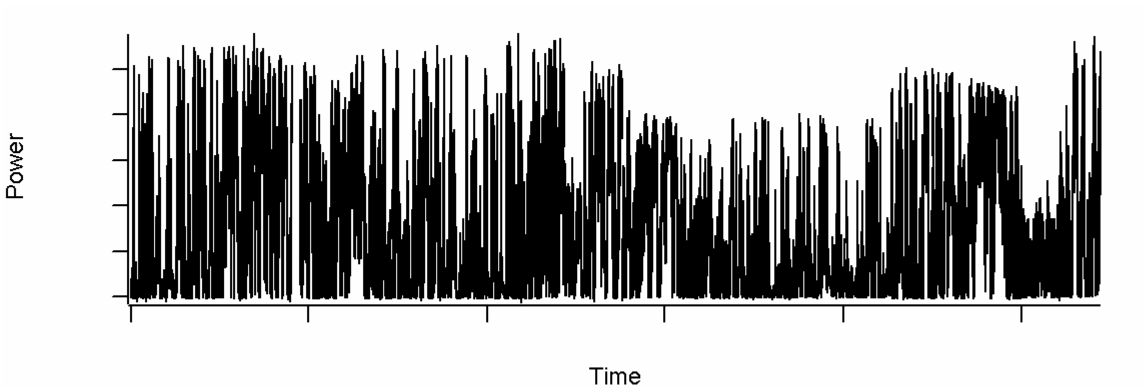

Figure 5.11. Generation of the Tehachapi wind aggregate in the CalSO multi-year dataset. January 2002 to September 2004.

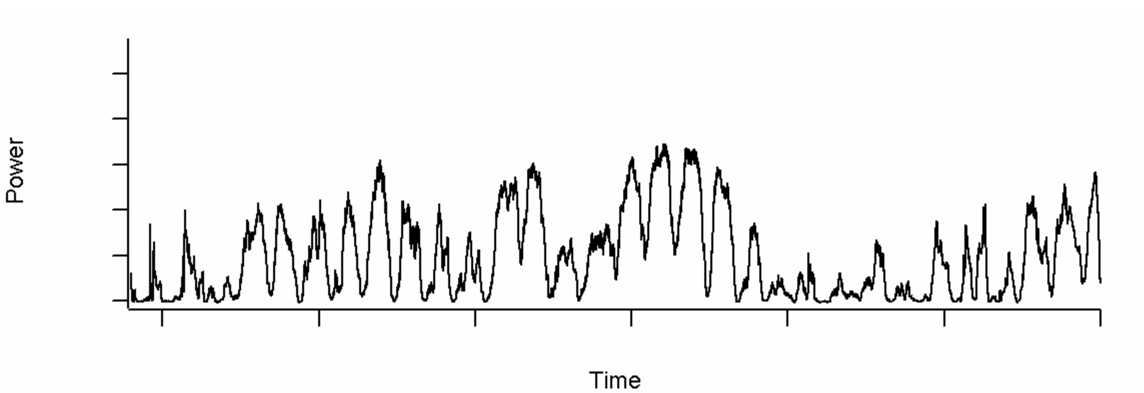

Figure 5.12. Generation of the Tehachapi wind aggregate in the CalSO multi-year dataset. One month in summer of 2003.

5.2.5 SCE DATASET

Southern California Edison provided an hourly dataset of aggregated renewable generation data that spanned 2002 to 2004. The data was recorded by SCE's revenue quality metering system, their highest quality data system. It did not exhibit the issues encountered with the CaISO multi-year dataset and was used directly in the capacity credit and load following analyses and to review and correct the one minute data in the CaISO multi-year dataset for the regulation analysis. The data was released through a confidentiality agreement. The items in the dataset are listed below.

Table 5.5. SCE dataset.

Data item	Reported nameplate capacity, MW			Annual peak, MW		
	2002	2003	2004	2002	2003	2004
Biofuel generation aggregate	207	205	210	153	148	148
Municipal waste generation aggregate	47	47	47	45	45	44
Geothermal (outside service territory) generation aggregate	434	634	634	426	631	631
Geothermal (within service territory) generation aggregate	318	318	318	351	349	348
Small hydro generation aggregate	96	96	96	56	62	60
Solar generation aggregate	379	379	379	407	463	401
Wind generation aggregate, San Gorgonio	357	362	362	325	317	332
Wind generation aggregate, Tehachapi	652	659	659	584	568	574

5.2.6 PG&E DATASET

Pacific Gas and Electric also provided an hourly dataset of aggregated renewable generation data. The data spanned 2002 to 2004 and was extracted from PG&E's settlement tables, their highest quality data system. Some minor issues were encountered with the treatment of timestamps and Daylight Saving Time, but the data did not exhibit any of the more serious issues encountered with the CaISO multi-year dataset. As with the SCE data, the PG&E data was used directly in the capacity credit and load following analyses and to review and correct the one minute data in the CaISO multi-year dataset for the regulation analysis. The data was released through a confidentiality agreement. The items in the dataset are listed below.

Table 5.6. PG&E dataset.

Data item	Average reported nameplate capacity, MW			Annual peak, MW		
	2002	2003	2004	2002	2003	2004
Biofuel generation aggregate	955	1007	1018	427	446	467
Geothermal generation aggregate	679	679	679	489	463	462
Wind generation aggregate	306	414	414	139	241	241

There appear to be inconsistencies in the reported nameplate capacities of the aggregates. In many cases, the nameplate capacity provided by PG&E is significantly higher than the annual peak generation. While a generator would not necessarily be expected to constantly produce power at its nameplate capacity, it should at least occasionally approach it. As shown above, the nameplate capacity is as much as 126% greater than the annual peak generation value. This discrepancy only affects the capacity credit analysis and capacity credit results are presented using both the reported nameplate capacity and the annual peak generation. This is discussed further in Section 5.3.5.

5.3 Data Issues

A variety of data issues were encountered in the various datasets used in the analysis. They are discussed below along with the methods used to address them.

5.3.1 CONFIDENTIALITY

Although the need to preserve the confidentiality of much of the study data is recognized, data confidentiality significantly impeded the study at several occasions. Establishing the initial data nondisclosure agreement (NDA) with CaISO was a very lengthy process. The experience garnered from the completion of this first NDA was valuable later in the study, as new study participants were able to receive draft NDAs from CaISO quickly.

Some other NDA processes were not as successful. In particular, SCE and NREL were unable to reach a confidentiality agreement even after numerous exchanges between their lawyers. Consequently, another analyst had to be trained to perform the capacity credit analysis, delaying the progress of the study.

Even with NDAs in place, the data released was aggregated because of concerns about the proprietary nature of power generation data from individual plants. Data aggregation aggravated data issues in the CaISO one-year and multi-year datasets, as discussed below. Later in the study, CaISO made a notable effort to allow the study analysts to view non-aggregated data while on-site at the CaISO offices; again, this is discussed further below.

5.3.2 MANAGEABILITY

The sheer size of the data is a problem, particularly with one minute data as in the CaISO one-year and multi-year datasets. To assemble the renewable aggregates, CaISO had to extract more than eighty pieces of raw data, each with 525,600 values per year. Even with automated retrieval scripts, extensive computer time was required to query such a large volume, especially in the case of the three year dataset. Because the disk space requirement for storing all of the individual data items was considered to be too great, CaISO calculated aggregated values as the individual data items were being retrieved; only the aggregated value was stored and individual data values were immediately discarded. The lack of ready availability of non-aggregated data later hindered the data review process.

Performing the data review and error checks for so much data was also a time intensive process. Because of the difficulties introduced by aggregation, the effectiveness of automated data checks was limited and all of the CaISO one minute data required manually review. The errors discovered in the one-year and multi-year datasets revealed an underlying problem. Because much of CaISO's data is stored automatically and is never used for operations or in any other way, it does not undergo any inspection except for generic automated tests by the PI system. Much of the data is therefore recorded without any verification of the quality of the data or the actual recording process.

5.3.3 LOSSY COMPRESSION

CaISO's PI system records over 180,000 pieces of data, some sampled many times a minute. To store so much data, a lossy compression scheme is used. Lossless compression uses algorithms that reduce the size of data while maintaining complete fidelity; when the data is uncompressed, it is exactly identical to what it was before compression was applied. Lossy compression sacrifices some accuracy for large improvements in size reduction; when the data is uncompressed, it is not exactly identical to what it was originally, but the changes should be negligible. The PI system uses the "Swinging Door" algorithm, a lossy scheme with configurable settings that trade off data fidelity and size. Ideally, information removed by compression is insignificant. However, the regulation analysis tracks even small fluctuations over short time periods. Data compressed without consideration for this type of calculation may affect the analysis when regulation impacts are small. Inspection of the data and regulation results suggests that the effects of compression might be significant only at impact levels when the regulation cost is negligible anyway.

5.3.4 TIMESTAMPS AND DAYLIGHT SAVING TIME

While outwardly trivial, timestamps and Daylight Saving Time must be handled carefully to ensure that datasets from different sources are correctly aligned. Data can be stamped with the time at the beginning, end, or middle of its sampling interval. Daylight

Saving Time is also treated in a variety of ways across datasets. Most often, data streams followed the active time standard, automatically shifting between Pacific Standard Time and Pacific Daylight Time in April and October as necessary. OASIS uses a particularly interesting method, keeping a twenty-fifth hour of data every day. Values of the “lost” or “extra” hour which occur when changing to or from Daylight Saving Time are stored in the twenty-fifth hour records.

Having a clearly defined timestamping convention for each dataset is obviously preferable. Absent that, datasets can be aligned by comparing similar data from different sources. For example, when comparing the CaISO multi-year dataset with the PG&E dataset, it was discovered that the raw PG&E data uses a lagging timestamp (e.g., the data for 12:00-13:00 has a timestamp of 13:00) whereas the CaISO and SCE data use leading timestamps (e.g., the data from 12:00-13:00 has a timestamp of 12:00). The CaISO and PG&E comparison also revealed an inconsistency in the handling of Daylight Saving Time in the PG&E wind data.

5.3.5 NAMEPLATE CAPACITY

Some inconsistencies appear in the values of the total nameplate capacity of the CaISO and PG&E generation aggregates. A discrepancy in the CaISO nameplate capacities was discovered while investigating a separate data issue. In this case, the power output of a wind plant exceeded its nameplate capacity by an order of magnitude. Because of this in combination with the elevated floor issue discussed further below, the CaISO data was replaced with IOU data in the capacity credit calculation, the only analysis affected by the values of the nameplate capacity. Table 5.7 compares the annual hourly power peaks with the reported nameplate capacities of some of the IOU generation aggregates.

Table 5.7. Comparison of reported nameplate capacities and annual peak generation of selected generation aggregates from the PG&E and SCE datasets.

Data item	2002			2003			2004		
	Name plate (MW)	Annual peak (MW)	% Diff	Name plate (MW)	Annual peak (MW)	% Diff	Name plate (MW)	Annual peak (MW)	% Diff
Biofuel generation aggregate, PG&E	955	427	124%	1007	446	126%	1018	467	118%
Geothermal generation aggregate, PG&E	679	489	39%	679	463	47%	679	462	47%
Wind generation aggregate, PG&E	306	139	120%	414	241	72%	414	241	72%
Geothermal generation aggregate, SCE, within service territory	318	351	-9%	318	349	-9%	318	348	-9%
Solar generation aggregate, SCE	379	407	-7%	379	463	-18%	379	401	-6%
Wind generation aggregate, SCE, San Gorgonio	357	325	10%	362	317	14%	362	332	9%
Wind generation aggregate, SCE, San Gorgonio	652	584	12%	659	568	16%	659	574	15%

In almost all cases, the nameplate capacity exceeds the annual generation peak. The SCE solar and geothermal (within service territory) aggregates are the only exception, with annual generation peaks exceeding nameplate capacities in all three years. In the case of solar, this is most likely because the nameplate capacity excludes the solar plants' auxiliary gas generators while the generation data includes their output.

While generators are not necessarily expected to operate at their nameplate capacity consistently, it is reasonable to expect them to at least approach that value occasionally. In the table above, the PG&E nameplate capacities are up to 126% greater than their corresponding generation peaks, whereas the difference is at most 16% in the SCE data used in the analysis. It is possible that the PG&E nameplate capacities have not been updated as plant operations have changed or individual generators in plants have been retired.

To address this issue, two sets of capacity credit results were presented: one based on the reported nameplate capacity and one on the annual generation peaks.

5.3.6 SPIKES AND DROPOUTS

Data spikes and dropouts appear throughout the CaISO one-year and multi-year datasets. These types of errors commonly occur in any measured data series and can be caused by faults in instrumentation or telemetry. These errors usually occur over very short periods of time, so they are most apparent in the faster sampled datasets (the one minute CaISO one-year and multi-year datasets) but less so in others where short periods of errors are suppressed by an hourly average. There is also some tradeoff between data quality and sampling rate, so again, these errors appear more in the faster sampled data.

In some cases, the spikes and dropouts are large enough and sharp enough that they can be easily detected both by visual inspection of the data and by simple automated checks of the first derivative of the generation data. Figure 5.13 shows an example of a large dropout in the CaISO dataset that is easily identifiable.

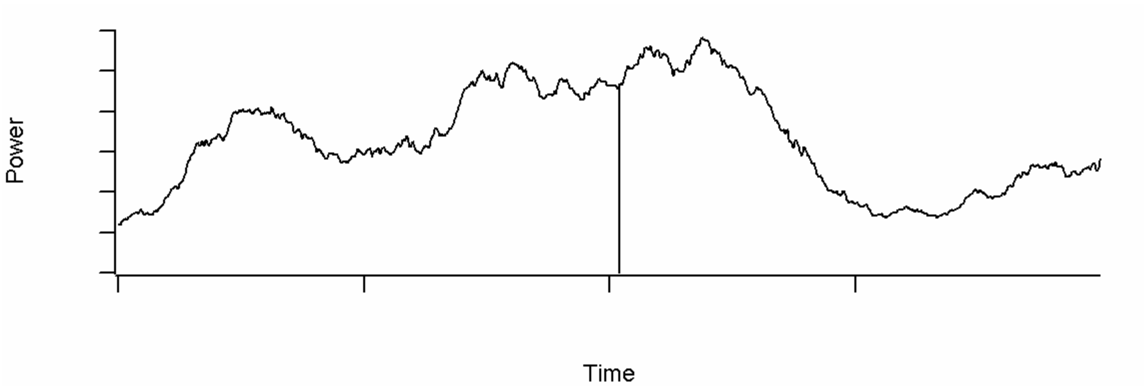

Figure 5.13. One day from the CaISO multi-year dataset showing a large dropout.

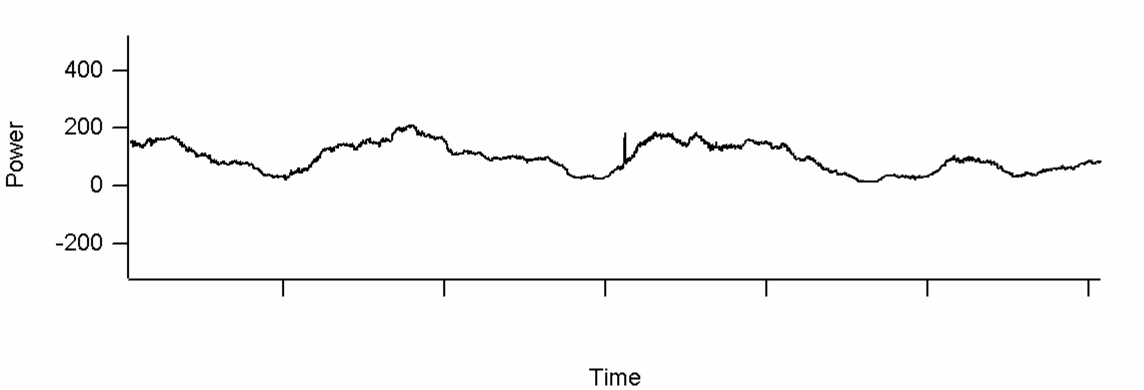

Figure 5.14. Three days from the CaISO multi-year dataset showing a data spike.

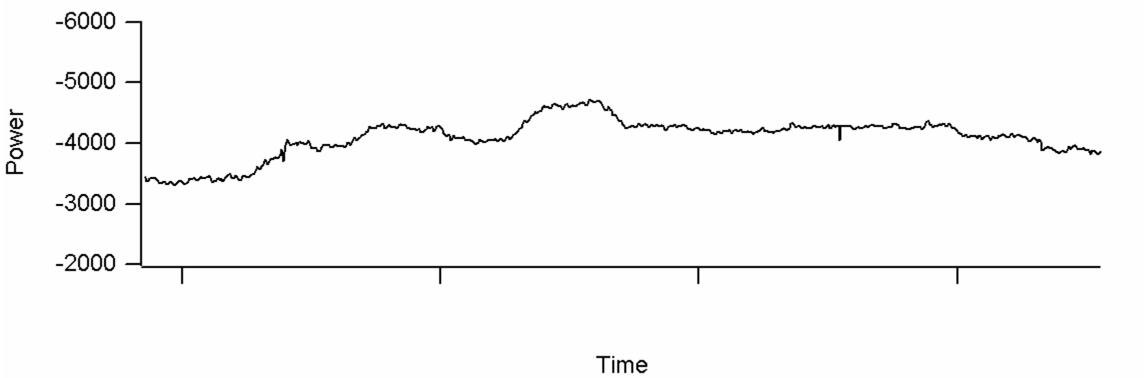

Figure 5.15. A twelve hour period from the CaISO multi-year dataset showing a small dropout.

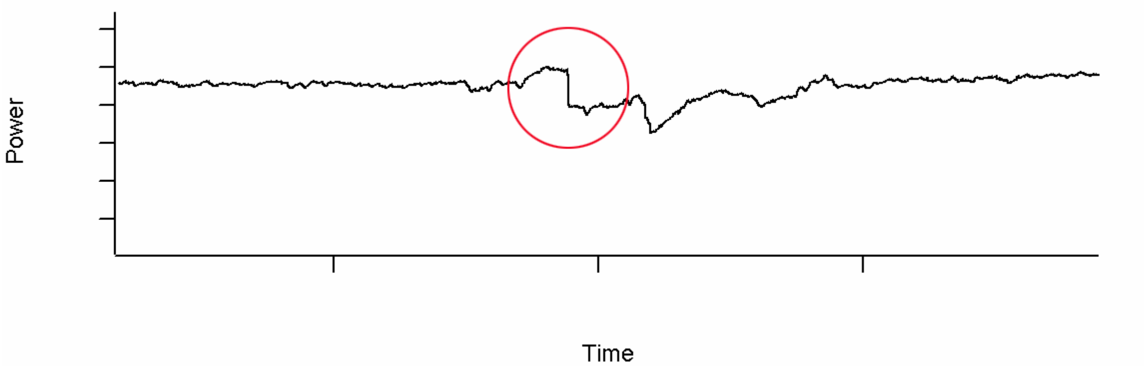

Figure 5.16. One day from the CaISO multi-year dataset showing a sharp 40 MW drop suspected to be a partial dropout in the data aggregate.

Spikes and dropouts, however, can be much more difficult if not impossible to detect when data is aggregated. As stated above, when the generation data stream of a single plant sharply rises or drops to zero at a physically impossible rate, it is generally easy to detect. If that generation data is aggregated with several other pieces of data of comparable magnitude or with a few other pieces of much larger magnitude, then it can be difficult or impossible to distinguish between a data error and a real ramp as shown in the 40 MW drop in Figure 5.16. The aggregated generation data from the CaISO one-year and multi-year datasets was all manually reviewed by visual inspection to identify spikes and dropouts. As demonstrated further below, the IOU data was used as a basis of comparison in the review process. Because the IOU data is hourly, it cannot be used to find every spike and dropout; however, the majority of significant ones can be found.

5.3.7 ELEVATED DATA FLOORS

Several of the data aggregates in the CaISO multi-year dataset exhibited periods where the data did not return to zero for prolonged periods as expected. When this occurred, the data appeared to be offset by either a constant value or a linear ramp as shown in Figure 5.17 and Figure 5.18.

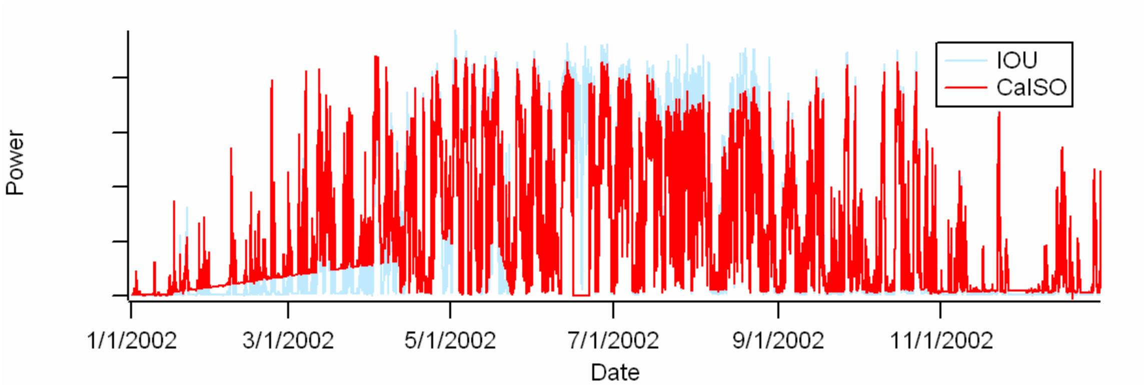

Figure 5.17. One year of data showing artificially elevated data floors. The red trace is from the CalSO multi-year dataset. The light blue trace behind it is corresponding IOU data.

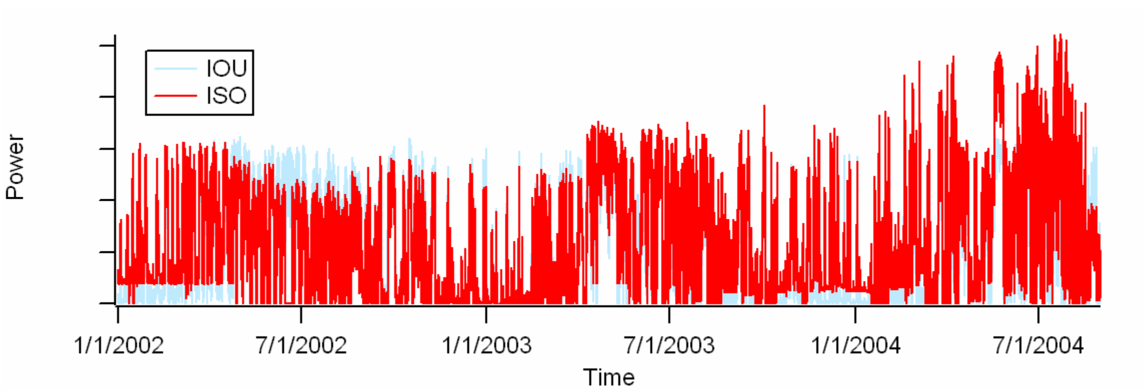

Figure 5.18. Almost three years of data showing artificially elevated data floors. The red trace is from the CalSO multi-year dataset. The light blue trace behind it is corresponding IOU data.

Figure 5.17 shows an elevated floor ramping up to approximately 70 MW over a four month period at the beginning of the year. The IOU data – the light blue trace behind the red – reveals that it is artificial. Two other ramped floors appear in the following months and a small constant offset occurs at the end of the year. In mid June, there is also a data dropout. In Figure 5.18, a number of offset floors appear again. Note that in May of 2003, there is a period in which the floor appears to be artificially elevated, but the IOU data reveals that this was real.

Further investigation revealed that the elevated floors were an artifact of the PI system's data compression routine. As introduced above, the PI system uses a lossy compression scheme that only stores new data points after a prescribed threshold of change has occurred. When the data is retrieved, the PI system uses an interpolation routine to fill in the data values between the stored points. An elevated data floor occurs when one of an aggregate's constituent data streams drops out and the dropout is not recognized as a data error. The PI system records valid data at the beginning and end of the dropout with the assumption that data was not recorded for the intervening period because of normal data compression. When the data is retrieved, the PI system fills the dropout period with interpolated values. If the values at the beginning and end of the dropout period are approximately the same, then a constant value is inserted; if not, then a ramp is inserted. This is shown in Figure 5.19 in which the entire aggregate dropped out, resulting in a perfect ramp in the data from 180 MW to 0 MW. In most cases, only part of the aggregate drops out and the constant offset or ramp "elevates" the rest of the data in the aggregate.

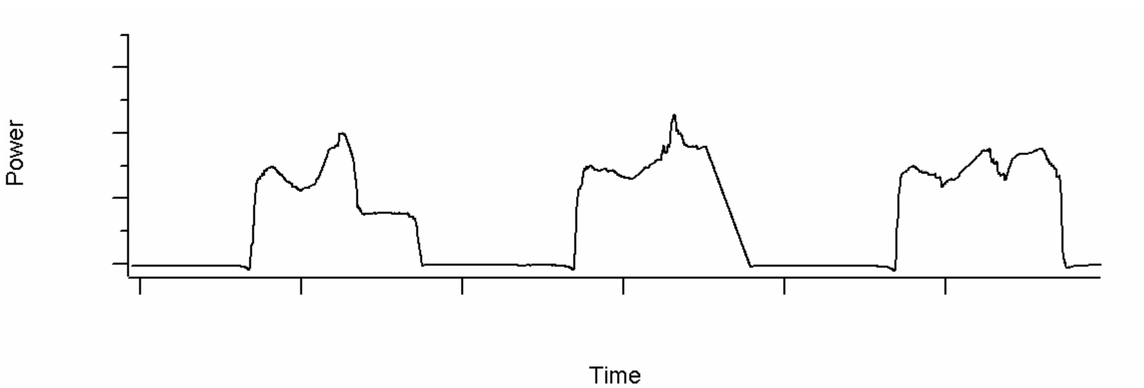

Figure 5.19. Three days from the CalSO multi-year dataset showing an occurrence of the dropout/interpolation error.

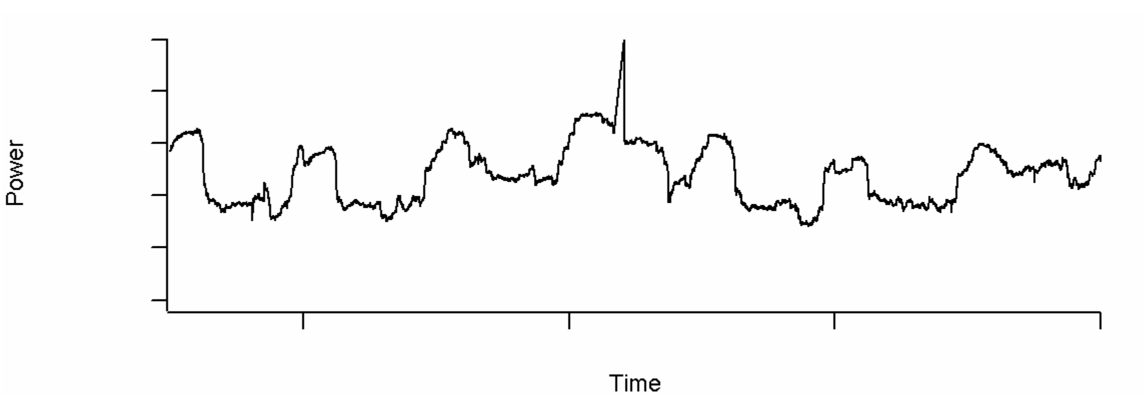

Figure 5.20. One week from the CalSO multi-year dataset showing a dropout/interpolation error immediately followed by a data spike.

Figure 5.20 shows one more occurrence of the dropout/interpolation error. In this case, a data spike immediately following the dropout period makes both errors easily visible.

As shown in the figures above, this type of error sometimes can be found through visual inspection of the data. However, there are cases in which it is not so apparent. Data floors elevated by a constant value can be hard to find in data with little variability; the offset cannot be readily distinguished from normal generating behavior. In this case, the erroneous interpolation could be hard to detect even when looking at the individual offending data stream uncoupled from the aggregate.

The dropout/interpolation error can also occur without any sort of elevated data floor. If one of the generation components of an aggregate experiences a dropout period which begins and ends at small or zero values, then the output of that component will be held near or at zero. The aggregated generation value during this period excludes the contribution of the component experiencing the dropout and the aggregate appears to generate less than it actually is.

The expansion of the data aggregates introduced this error to the multi-year dataset. Including more PI tags increased the capacity of the aggregates, but also increased the chances for this error to appear in the dataset. CaISO reconfigured the PI system in January 2004 so that interpolation would be applied only to dropouts of 7.5 minutes or less. This window was further reduced in September 2004 to one minute. This reduction should prevent the error from occurring in subsequent data.

Several methods were considered to address this issue as described in the following section. Ultimately, IOU data was used as a basis of comparison to identify and correct the elevated floors and other occurrences of the dropout/interpolation error.

5.3.8 AGGREGATION AND DATASET COMPARISON

With accurate data, the aggregation of generation data can reflect the real-world aggregated behavior of generators acting simultaneously in a system. However, in general, data problems in the individual components of an aggregate are obscured by aggregation, making inaccuracies in a data aggregate hard to identify. The spikes, dropouts, and elevated floor/interpolation errors in the CaISO multi-year dataset discussed above are all more difficult or impossible to find in aggregated data.

Several methods were considered to address the data quality issues in the aggregated generation data in the CaISO multi-year dataset:

- Use of the trustworthy aggregated capacity
- Review and correction of the individual data components of the aggregates
- Use of data from other sources

As described in Section 5.2.4, the trustworthy aggregated capacity was provided by CaISO in the multi-year dataset to identify periods in the data with problems. For a given aggregate at a given time, it is the summed capacity of the generators to have good data. Whether or not a data point is considered good or not is determined by an algorithm developed by CaISO using data quality tags and timestamps recorded by the PI system. Other solutions to the data quality issues were pursued for two reasons. First, the capacity credit analysis requires a dataset with consistent capacity over one year. Even if periods with bad data were identified and the amount of capacity reporting good data was accurately known, the data requirements for the capacity credit analysis would not be met. Second, the trustworthy aggregated capacity did not always accurately reflect the amount of capacity reporting good data. Errors were found in the data that were not identified by the algorithm used to calculate the trustworthy aggregated capacity. In later comparisons of the CaISO multi-year dataset with IOU datasets, it was discovered that the trustworthy aggregated capacity also flagged a large number of false positives.

Inspecting the individual data components of the aggregates was the first solution considered to address the data quality issues. There were two problems with this approach, related to the confidentiality and manageability issues discussed above. First, the non-aggregated data was not accessible. CaISO eventually determined that the data could be viewed while on-site at their offices and they made arrangements to do so with considerable effort. Second, CaISO and the analysis team agreed that this would be a very large, time intensive effort with unknown complications. The amount of generation data requiring review would increase by an order of magnitude. Complicating matters, many of the constituent components of the data aggregates are aggregates of several plants themselves, received and stored by CaISO only in aggregated form. It was presumed that further inspection of these sub-aggregates would reveal data quality issues of their own, prolonging the overall data review. Inspection of the individual data components was pursued simultaneously with the IOU data comparison, described below. When the IOU data was determined to be sufficient to address the data quality issues, efforts were focused there and the inspection of the individual data components was abandoned.

Additional data sources were ultimately used to resolve the data quality problems. PG&E and SCE were able to provide verified, high quality data that was used directly in the capacity credit and load following analyses. However, neither had data sampled fast enough for the regulation analysis. While the aggregates in the IOU and CaISO multi-year datasets were not exactly the same, they were identical enough that the IOU data could be used to find errors in the CaISO multi-year dataset. Programs were developed to automate parts of the comparison and to assist with the manual review of the datasets.

Figure 5.21.Screenshot of one of the programs developed to process the CalSO multi-year dataset using IOU data as a basis of comparison.

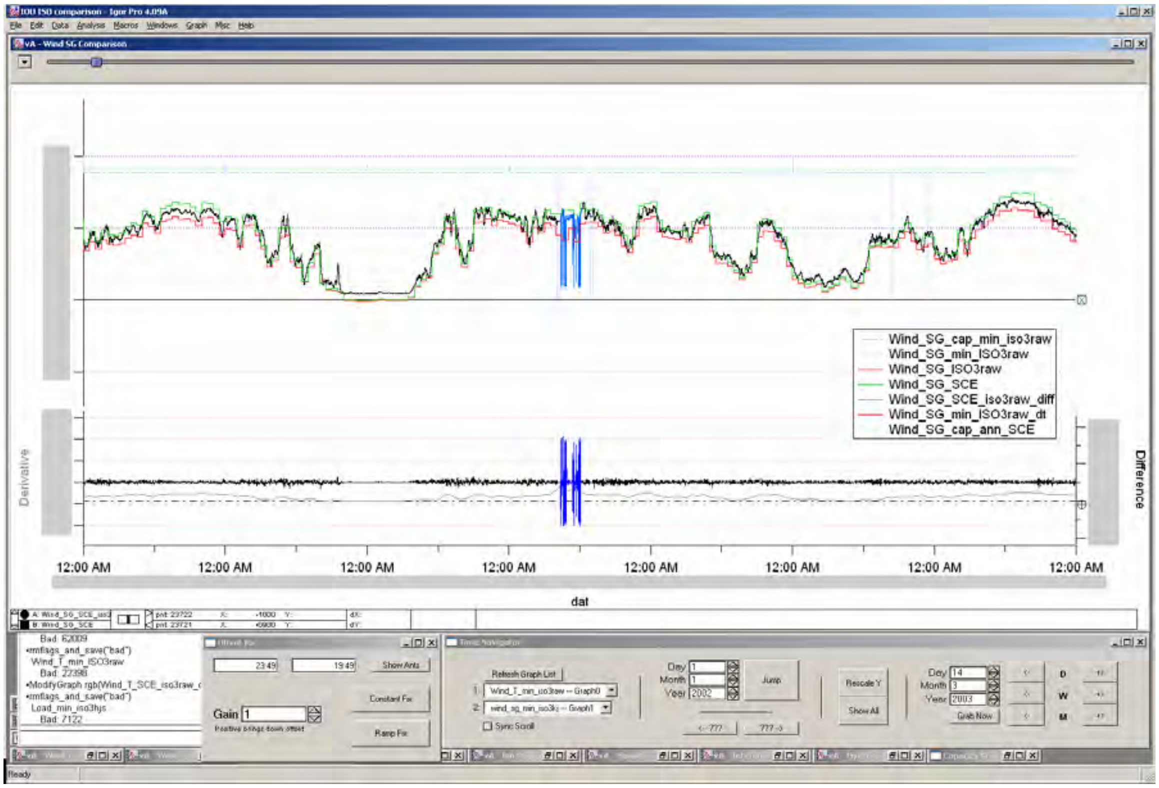

Figure 5.22.Screenshot of one of the programs developed to process the CalSO multi-year dataset using IOU data as a basis of comparison.

Figure 5.21 and Figure 5.22 show some screenshots from the data comparison process. The program overlays the following data for a given aggregate:

- CalSO one-minute generation
- Hourly average of the CalSO one-minute generation
- IOU hourly generation
- CalSO trustworthy aggregated capacity
- Difference between IOU and CalSO hourly generation
- Generation ramp rate calculated from the CalSO one-minute generation data

Through manual comparison of the datasets with automated algorithms to identify suspect data, spikes and dropouts were removed and elevated data floors were corrected in the CalSO multi-year dataset.

5.3.9 DATA GAPS

All of the datasets begin on January 1, 2002 and run through most, if not all, of 2004. Some are missing the end of 2004 and data is missing at various other points where data quality or extraction errors occurred. Data gaps affect the capacity credit, regulation, and load following analyses in different ways, as described below.

The capacity credit analysis requires a complete year of synchronized data. Because 2004 data is not complete, the third year of analysis is not the full calendar year of 2004, but one year from mid-September 2003 to mid-September 2004. Some of the renewable generation data had a full three calendar years, but September 22, 2004 is the last point in time when data was available in all the datasets. Small gaps exist in the one minute CaISO multi-year load, hydro, and interchange data where bad data was found and removed. Because only hourly averages of the one minute data are used, the gaps are in almost all cases short enough that they do not affect the analysis. A larger gap exists in the hydro data from February 14, 2002 to February 26, 2002. Because this is a low risk period with a small contribution to the overall annual risk, the gap was simply patched with scaled data from another period. The IOU renewable generation data used in the analysis did not have any gaps.

The regulation analysis is tolerant of data gaps as long as enough hours are included to be representative of the analysis period. Large gaps from mid-February to mid-September 2002 appear in the biomass and solar data from the CaISO multi-year dataset. Because results from the 2002 analysis of the CaISO one-year dataset were available for comparison, the 2002 analysis of the CaISO multi-year dataset proceeded with the biomass and solar data gaps. The 2002 biomass and solar regulation analyses were conducted normally; the 2002 regulation analyses of the other renewables excluded biomass and solar from their calculation of the total system compensation requirement (Equation 3.8) to keep a relatively intact year long dataset. The results matched well with the 2002 results from the one-year dataset. The 2004 analysis covers January 1, 2004 to September 19, 2004.

Like the regulation analysis, the load following analysis is tolerant of data gaps as long as enough hours are included to be representative of the analysis period. The analysis used renewable generation data from the IOU datasets and hourly values of actual, forecasted, and scheduled load from OASIS. There were almost no gaps in these datasets except for the Northern California wind aggregate from PG&E, which is missing December 2004.

6 RECOMMENDATIONS

The Phase III report made several recommendations about the implementation of integration cost analysis. Based on experiences from the multi-year analysis, the following additional recommendations pertaining to data reporting/collection and an Integration Cost Analyst (ICA) are proposed.

6.1 Data Reporting and Collection

The majority of time and effort required for the multi-year analysis was dedicated to data collection and processing. The actual calculations and review of the results were relatively straightforward. Specific recommendations are therefore made for the handling of data for future integration cost analysis.

In Phase III of the study, it was proposed that data collection should be performed by an Integration Cost Analyst, a California Energy Commission or CPUC staff tasked with performing and reporting on regular integration cost analysis. Given the complex data quality issues described in Section 5.3 and the need for similar data in other recent and current studies such as the Energy Commission’s Strategic Value Analysis and Intermittency Analysis Project, it is now recommended that data handling and integration cost analysis be separated into two distinct tasks. A data handling entity would be responsible for collecting, reviewing, storing, and providing data for integration cost analysis and, possibly, associated data for other studies. In Phase III, it was assumed that data collection and processing was essentially an accounting function which would be highly automated. While this eventually may become true, given the data issues described in this report, data handling is more appropriately an engineering task. The data handling entity would have to meet the following requirements and perform the following duties:

- Satisfy confidentiality requirements of CaISO, IOUs, and other sources to access data.
- Provide a database that securely stores data and that can be easily queried for both manual and automated data input and retrieval.
- Coordinate with CaISO, IOUs, and other sources to receive data on a frequent, regular basis; a one month basis is recommended. Jointly develop a reporting standard with the data sources for incoming data and, as necessary, tools to process various data types and formats. Also, jointly develop an automated reporting system so that data is transferred from the sources to the data handling entity automatically. Update data requests as necessary as new generators come online and other changes occur.
- Review and verify the quality of incoming data and flag and/or correct bad data.

- Coordinate with CaISO, IOUs, and other sources as necessary to ensure that the quality of data they are collecting and recording is sufficient for the intended analyses. As ongoing integration cost calculation is presumed for the future, this process should begin immediately.

- Coordinate with the Integration Cost Analyst to ensure that the required data is collected with sufficient quality and provided to the ICA on a frequent, regular basis; again, a one month basis is recommended. Jointly develop a reporting standard with the Integration Cost Analyst and an automated system for transfer of data from the data handling entity to the ICA.

One of the key aspects of the proposed data handling process is that the assurance of data quality is a shared responsibility between the data sources (CaISO, IOUs, etc.), the data handling entity, and the Integration Cost Analyst. The task otherwise becomes disproportionately difficult to manage and complete.

It is also important to collect and review data on a frequent and regular basis. Many of the difficulties encountered with the processing of the datasets for the multi-year analysis were the result of working with such a large, lumped amount of data at once. As originally proposed in Phase III, it is recommended that data be documented monthly in arrears for the previous month. Processing data on a frequent basis not only keeps the task more manageable, but allows errors and issues to be identified and corrected before they propagate into a larger amount of data over an extended period. Automated data reporting would simplify the collection process, but the data review will always include some manual inspection.

6.2 Integration Cost Analyst

An Integration Cost Analyst (ICA) was introduced in Phase III and is recommended again with some revisions to the original description of qualifications and responsibilities. The function of the ICA is to perform regular analysis and reporting of integration costs. It is proposed that the California Energy Commission or CPUC designate one or more staff to assume this role. Specifically, the ICA would have to meet the following requirements and perform the following duties:

- Satisfy confidentiality requirements of CaISO, IOUs, and other sources to access data.

- Coordinate with the data handling entity previously described and, as necessary, the various data sources to ensure that all required data is of sufficient quality and is received on a frequent, regular basis in a consistent format. Again, it is recommended that data be received on a monthly basis.

- Review incoming data as it is received to verify data quality.

- Annually perform integration cost analysis.

- Prepare annual reports documenting the results of the integration cost analysis.

Assuming the availability of good data, the calculations involved in integration cost analysis are relatively straightforward and can be highly automated. Once procedures are established and refined, it is estimated that the ICA will require approximately one to two days per month to perform data handling tasks and approximately two additional weeks each year to conduct the integration cost calculations, perform an analysis of the results, and generate a report.

REFERENCE

[1] Sher et al, California Senate Bill 1078, chaptered 12 Sep 2002.

[2] Kirby, B., M. Milligan, Y. Makarov, D. Hawkins, K. Jackson, and H. Shiu, "California RPS Integration Cost Analysis – Phase I: One Year Analysis of Existing Resources", Consultant Report 500-03-108C, California, California Energy Commission, December 2003.

[3] Kirby, B., M. Milligan, Y. Makarov, D. Hawkins, J. Lovekin, K. Jackson, and H. Shiu, "California Renewables Portfolio Standard Renewable Generation Integration Cost Analysis – Phase III: Recommendations for Implementation", Consultant Report P500-04-054, California, California Energy Commission, July 2004.

[4] Piwko, Osborn, Gramlich, Jordan, Hawkins, and Porter, "Wind Energy Delivery Issues", *IEEE Power and Energy*, Vol. 3, No. 6, Nov/Dec 2005.

[5] U.S. Federal Energy Regulatory Commission, *Regional Transmission Organizations, Notice of Proposed Rulemaking*, Docket No. RM99-2-000, Washington, DC, May 13 1999.

[6] U.S. Federal Energy Regulatory Commission, *Promoting Wholesale Competition Through Open Access Non-Discriminatory Transmission Services by Public Utilities; Recovery of Stranded Costs by Public Utilities and Transmitting Utilities, Final Rule*, Docket Nos. RM95-8-000 and RM94-7-001, Order No. 888, Washington, DC, April 24 1996.

[7] Hirst, E. and B. Kirby, "Unbundling Generation and Transmission Services for Competitive Electricity Markets: Examining Ancillary Services", NRRI 98-05, National Regulatory Research Institute, Columbus, OH, January 1998.

[8] North American Electric Reliability Council, *NERC Operating Manual*, Princeton, NJ, November 2002.

[9] E. Hirst and B. Kirby, "Measuring Generator Performance in Providing the Regulation and Load-Following Ancillary Services", ORNL TM/2000-383, Oak Ridge National Laboratory, Oak Ridge, TN, December 2000.

[10] Electronic Wind Performance Reporting System, http://wprs.ucdavis.edu/, accessed in October 2005.

[11] Kirby, B. and M. Milligan, "A Method and Case Study for Estimating The Ramping Capability of a Control Area or Balancing Authority and Implications for Moderate or High Wind Penetration", AWEA WindPower 2005, Denver, CO, May 2005.

APPENDIX A: CONTROL PERFORMANCE STANDARDS[1]

The electrical power system operated by the *California Independent System Operator* (CaISO) is called its *control-area*. Power plants, or *generators*, located throughout the state are managed in real-time to meet the demands, or *loads,* of electricity customers. Because electricity is a real-time product in which loads and generation fluctuate and cannot be perfectly predicted, control-area operators, or *dispatchers*, must constantly adjust generation to meet load. CaISO manages electrical *energy*, generating *capacity*, and other *ancillary services* that are used to maintain control and reliability of the California utility grid.

The CaISO must manage its generators to compensate for the real-time variations between actual generation and actual load in the electric system. The *North American Electric Reliability Council* (NERC) recognizes the *area control error* (ACE) as a primary metric used to assess the performance of the control operator. Each control area seeks to minimize its effects on the neighboring control areas to which it maintains an *interconnection*. Errors incurred because of generation, load or schedule variations or because of jointly owned units, contracts for regulation service, or the use of dynamic schedules must be kept within the control area and not passed to the interconnection. The equation for ACE is:

$$ACE = (NI_A - NI_S) - 10ß\,(F_A - F_S) - I_{ME} \qquad \text{Equation A.1}$$

In this equation, NI_A accounts for all actual meter points that define the boundary of the control area and is the algebraic sum of flows on all tie lines. Likewise, NI_S accounts for all scheduled tie flows of the control area. The combination of the two (NI_A - NI_S) represents the ACE associated with meeting schedules and if used by itself for control would be referred to as flat tie line regulation.

The second part of the equation, 10ß (F_A - F_S), is a function of frequency. The 10ß represents a control area's frequency bias (ß's sign is negative) where ß is the actual frequency bias setting (MW/0.1 Hz) used by the control area and 10 converts the frequency setting to MW/Hz. F_A is the actual frequency and F_S is the scheduled frequency. F_S is normally 60 Hz but may be offset to effect manual time error corrections. I_{ME} is the meter error recognized as being the difference between the integrated hourly average of the net tie line instantaneous interchange MW (NI_A) and the hourly net interchange demand measurement (MWh). This term should normally be very small or zero.

The North American Electric Reliability Council *Control Performance Standards* (CPS) 1 and 2 set statistical limits on the allowable differences between one-minute averages of the control area's difference between aggregated generation and interchange schedules relative to load (i.e., ACE). CPS1 measures the relationship between the control area's ACE and its interconnection frequency on a one-minute average basis. CPS1 values are recorded every minute, but the metric is evaluated and reported annually. NERC

[1] North American Electric Reliability Council. *NERC Operating Manual*. Princeton, NJ, November 2002.

sets minimum CPS1 requirements that each control area must exceed each year. CPS2 is a monthly performance standard that sets control-area-specific limits on the maximum average ACE for every 10-minute period.

Neither CPS1 nor CPS2 require that the ISO maintain a zero value for ACE. Small imbalances are generally permissible, as are occasional large imbalances. Both CPS1 and CPS2 are statistical measures of imbalance, the first a yearly measure and the second a monthly measure. Also both CPS standards measure the aggregate performance of the control area, not the behavior of individual loads or generators. Control areas are permitted to exceed the CPS2 limit no more than 10% of the time. This means that a control area can average no more than 14.4 CPS2 violations per day during any month.

APPENDIX B: REGULATION ALLOCATION METHODOLOGY

This regulation impact allocation method[2] was developed by Oak Ridge National Laboratory to deal with nonconforming loads. It works equally well with uncontrolled generators that are not using either AGC or ADS. The methodology meets several desirable objectives:

- Recognize positive and negative correlations
- Be independent of subaggregations
- Be independent of order in which generators or loads are added to system
- Allow disaggregation of as many or few components as desired

The methodology has been used by a number of analysts to analyze the regulation impacts of loads, conventional generators that are not on AGC or ADS, and non-dispatchable renewable generators.

We can think of regulation as a vector and not just a magnitude. For example, start with load *A*. It might be a single house or an entire control area with a regulation impact of 8. Consider another load *B* with a regulation impact of 6 that we want to combine with *A*. If loads *A* and *B* are perfectly correlated positively, they add linearly, as shown in the top of Figure B.1. If the two loads are perfectly correlated negatively, their regulation impacts would add as shown in the middle of Figure B.1. Typically, loads are completely uncorrelated and the regulation requirement for the total is the square root of the sum of the squares, or 10 in this case (bottom of Figure B.1).

Multiple uncorrelated loads are always at 90 degrees to every other load. They are also at 90 degrees to the sum of all the other loads. This characteristic requires adding another dimension each time another load is added, which is difficult to visualize beyond three loads. Fortunately, the math is not any more complex. The fact that each new uncorrelated load is at 90 degrees to every other load and to the total of all the other loads is quite useful. The analysis of any number of multiple loads can always be broken down into a two-element problem, the single load and the rest of the system.

Return to the two-load example but consider the more general case where loads *A* and B are neither perfectly correlated nor perfectly uncorrelated. We may know the magnitude of *A* and the magnitude of *B*, but we do not know the magnitude of the total without measuring it directly (i.e., we do not know the *direction* of each vector). We can, however, measure the total regulation requirement and use this vector method to *allocate* the total requirement among the individual contributors.

We know the total regulation requirement because we meter it directly as the aggregated regulation requirement of the control area. We can know the regulation requirement of any load by metering it also. We can know the regulation requirement of the entire system less the single load we are interested in by calculating the difference

[2] Kirby, B. and E. Hirst, "Customer Specific Metrics for the Regulation and Load-Following Ancillary Services", ORNL/CON-474, Oak Ridge National Laboratory, Oak Ridge, TN, January 2000.

between the system load and the single load at every time step, separating regulation from load following, and taking the standard deviation of the difference signal. Knowing the magnitudes of the three regulation requirements, we can draw a vector diagram showing how they relate to each other (Figure B.2).

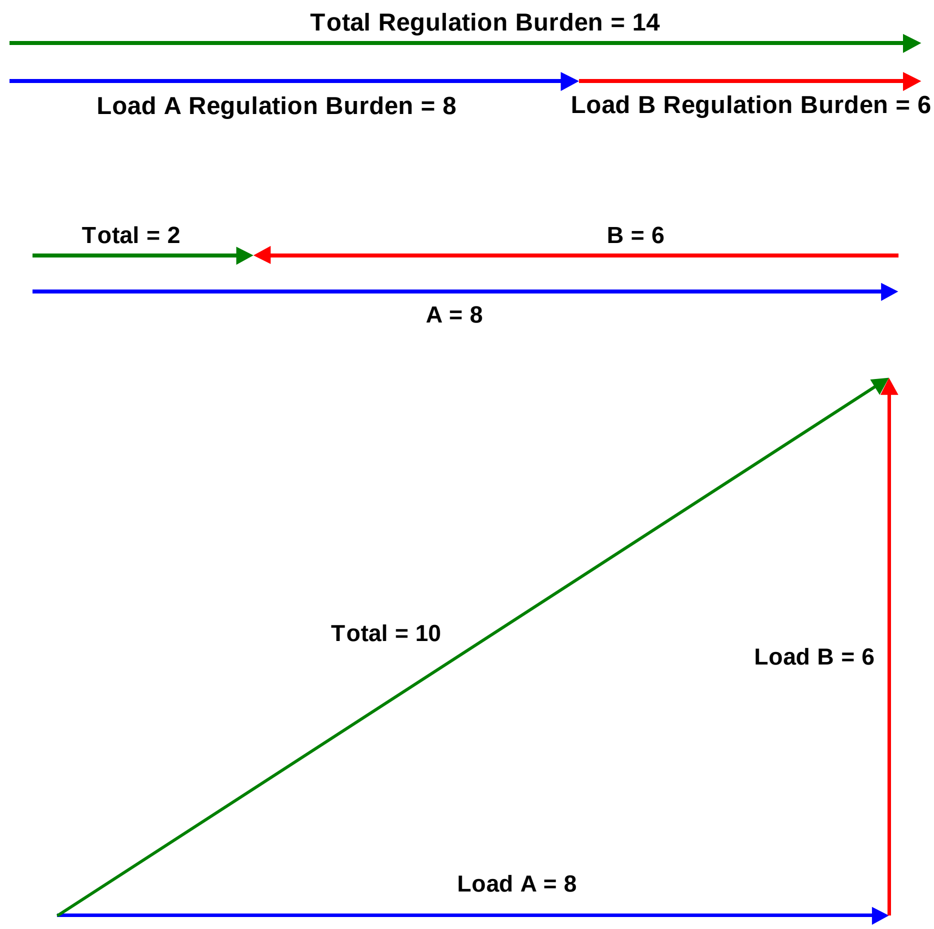

Figure B.1. The relationships among the regulation components (A and B) and the total if A and B are positively correlated (top), negatively correlated (middle), or uncorrelated (bottom).

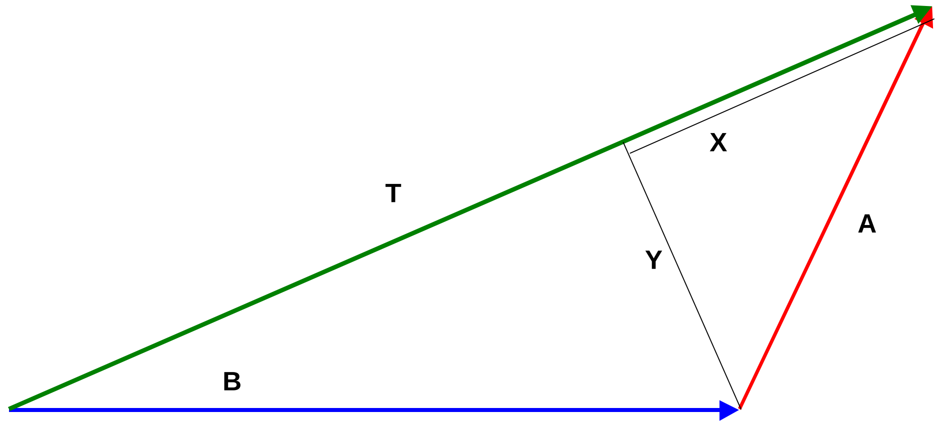

Figure B.2. The relationship among the regulation impacts of loads A and B and the total (T) when A and B are neither perfectly correlated nor perfectly uncorrelated.

How much of the total regulation requirement is the responsibility of load *A*? We can calculate the amount of A that is aligned with the total and the amount of *B* that is aligned with the total. We can do this geometrically (Figure B.2) or with a correlation analysis.

Y is perpendicular to the total regulation *T* (uncorrelated). *X* is aligned with *T* (correlated). *A*'s contribution to *T* is *X*. Knowing *A*, *B*, and *T*, we can calculate *X*. (We could also calculate *Y*, but there is no need to do so.) We can write two equations relating the lengths of the various elements:

$$A^2 = X^2 + Y^2 \qquad \text{Equation B.1}$$

$$B^2 = (T - X)^2 + Y^2 \qquad \text{Equation B.2}$$

Subtract Equation B.2 from Equation B.1 to get,

$$A^2 - B^2 = X^2 - (T - X)^2 + Y^2 - Y^2$$

$$A^2 - B^2 = X^2 - (T^2 - TX - TX + X^2) = 2TX - T^2$$

Solving for *X* (load *A*'s contribution to the total *T*) yields,

$$X = (A^2 - B^2 + T^2)/2T \qquad \text{Equation B.3}$$

We can decompose a collection of any number of loads into a two-load problem consisting of the load we are interested in and the rest of the system without that load (Figure B.3). We can solve Equation B.3 for as many individual loads as we wish. Variable *T* remains the total regulation requirement, variable *A* becomes each individual load's regulation requirement, and variable *B* becomes the regulation requirement of the total system *less* the specific load of interest.

This allocation method works well with any combination of individually metered loads and load profiling for the remaining loads. The load profiling can be as simple as making the usual assumption that the other loads' regulation requirements are proportional to their energy requirements. Or measurements of a sample set can be taken to determine the magnitude of their regulation impacts. This vector-allocation method is used to determine the regulation impact of each of the metered loads. The residual regulation impact is then allocated among the remaining loads, assuming they are perfectly uncorrelated.

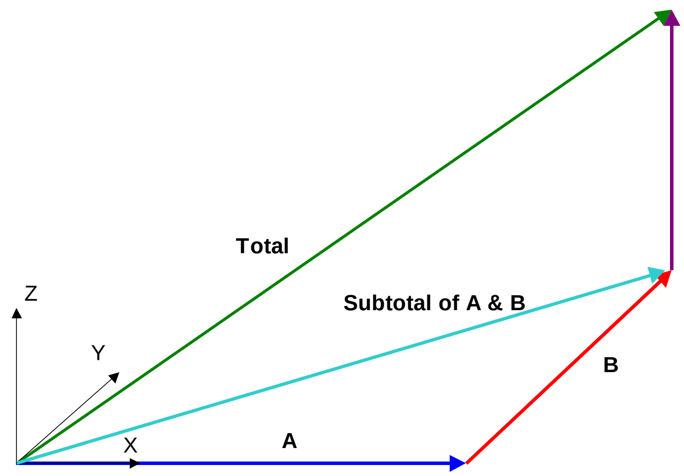

Figure B.3 Application of the vector-allocation method to the case with more than two loads.

APPENDIX C: COMMENTS FROM PACIFIC GAS AND ELECTRIC COMPANY

C.1. Received Comments

The comments on the next five pages were received following the release of a draft of this report in March 2006 and a public workshop discussing the findings of this study on 3 April 2006. The analysis team would like to again thank PG&E for their participation and insight throughout this study as well as for their consideration in preparing the comments herein.

Pacific Gas and Electric Company

Les Guliasi
Director
State Agency Relations

77 Beale Street, Room 2989
P.O. Box 77000, B29L
San Francisco, CA 94177-0001

415.973.6463
Fax: 415.973.9572

April 21, 2006

California Energy Commission
Attn: Cost of Integrating Renewables
[Pete Spaulding, MS-43]
1516 Ninth Street
Sacramento, CA 95814-5512

Re: Multi-year Integration Analysis

Dear Mr. Spaulding:

Attached are Pacific Gas and Electric Company's comments relating to the California Renewables Portfolio Standard Renewable Generation Integration Cost Analysis: Multi-Year Analysis Results and Recommendations. PG&E may provide additional comments soon. We look forward to continuing to participate in this effort.

If you have any questions, please do not hesitate to call me at (415) 973-6463.

Sincerely,

Les Guliasi/mld

Attachment

PG&E's Comments on the Draft Report: California Renewables Portfolio Standard (RPS) Renewable Generation Integration Cost Analysis Multi-year Intermittency Analysis

On April 3rd 2006, the California Energy Commission held a workshop in Sacramento to present the California Renewables Portfolio Standard (RPS) Renewable Generation Integration Cost Analysis Multi-year Intermittency Analysis (Multi-Year Analysis). PG&E recognizes the significant amount of work that went into this report and appreciates the opportunity to comment. Overall, the report is a good starting point for a framework for evaluating various cost components of integrating existing intermittent resources under historical operations and system conditions. The analysis validated earlier work but also highlighted areas that need further examination in the future. In the past, PG&E made general comments about the Phase 1 approach, and its methodology for determining Capacity Credit, Regulation and Load Following. Here, PG&E provides additional feedback about the methodology which it hopes will help identify and account for currently unaccounted or understated costs, and provides thoughts on the future application of this analysis and the proposed Integration Cost Analyst (ICA) position.

Capacity Credit

Effective Load Carrying Capability (ELCC) values in the Multi-Year Analysis validate earlier results. However, it is unclear how the ELCC, as a measure of impact to system reliability, would be used in Least Cost Best Fit (LCBF) and bid evaluation for intermittent resources like wind because it does not comport with Resource Adequacy accounting or procurement protocols. Capacity credit must be weighted to the value of capacity in the market, which typically changes seasonally. For example, capacity credits for wind production during the on-peak period in May and June are considerably higher than the July-September period when the value of that capacity is typically highest.

Furthermore, the capacity credit, using an ELCC methodology, does not capture the costs associated with back-up reserves that might be required by an Electric Service Provider (ESP) to balance load, generation and reserve requirements. For example, the 24% Northern California wind capacity credit (ELCC) reported in the Multi-Year Analysis is much higher than what the CAISO has historically observed for Northern California wind (< 5%) during system peaks[1]. Hence, the ESP may have to procure additional capacity to make up the difference. In the recent CPUC/CRS report "Achieving a 33% Renewable Energy Target", this extra procurement is referred to as "contingency reserves," the need for which is explained in terms of the characteristics of wind: "And while there is a reasonably good match between seasonal production profiles for California wind plants and seasonal peak demand, wind production often is quite low during the peak 50 to 100 hours per year."[2]

In order to account for these additional capacity credit costs, PG&E recommends that the CEC task the ICA with evaluating the feasibility of modifying the existing ELCC methodology to reflect current Resource Adequacy methodology so that monthly

[1] CAISO, 2005 Summer Operations Assessment, March 23, 2005, pg 19

[2] CRS/CPUC, Achieving a 33% Renewable Energy Target, November 1, 2005, pg 50

variations are captured, and also tasks the ICA with amending the methodology to account for the cost of "contingency reserves."

Regulation

PG&E believes the estimates of regulation costs in the Multi-Year Analysis are low based on two observations. First, the study only assesses costs and volumes of regulation from procurement of "reg-up" and "reg-down" (approximately 3% of system load) and does not capture additional regulation service that may be provided from units on AGC that are simultaneously providing spinning reserves. Second, the ISO's procurement of Ancillary Services (AS) in GWh changed significantly from year to year; the greater the volume of regulation the lower the average regulation prices. This suggests the incremental cost of regulation is sensitive to the amount of regulation the ISO procures (e.g. Year 2002 had the greatest volume of total regulation (Table 3.12) of 7220 GWh and lowest allocation of costs associated with Northern California Wind of $ -.24 MWh). Therefore, the concern is that incremental costs for regulation for an intermittent resource may get "lost" if ISO procures AS for other system needs. PG&E also highlights a finding that goes against conventional wisdom; the Multi-Year Analysis shows relatively similar regulation costs for solar and wind despite solar's less variable production and shape that is more complementary to load.

Finally, as more intermittent resources are added to the system and variability increases, there will be a greater impact and cost associated with integrating those resources into the system (see Historical Penetrations, below).

Load Following

The Multi-Year Analysis concludes that at current penetration levels, there is an ample depth in the resource stack to handle incremental energy. "At current penetration levels, the scheduling error of the renewables does not have a significant effect on the total energy requirements from the short-term market. The minimum scheduling bias reduced over the years but remained well over 200% greater than the load forecast error. This implies ample depth in the generator stack to handle incremental energy." However, this incorrect conclusion does not consider, from the ESP perspectives, the resources which have been pre-arranged to be available in the real-time markets. Day-ahead scheduling protocols have changed since the 2002-2004 period and now require ESP's to schedule generation to 95% of forecasted load on a day-ahead basis. Therefore, going forward, the remaining stack of generation committed to serve the 5% of remaining load may be "thin" resulting in potentially higher costs in hour-ahead markets and/or over/under scheduling of load following resources in day-ahead market.

Ramping Capability

The Multi-Year Analysis concludes "preliminary results shows that there is a very large amount of ramping capability." The Multi-year study did not assess ramping capability availability due to contractual constraints, during periods of high hydro conditions when ramping capability from hydro facilities is low and high maintenance periods (spring/fall).

Use of Historical Penetrations
In its Phase 3 report, the CEC proposed that the results of this study "will reveal the integration impacts of present generation in specified areas," and that "these results can act as a proxy for the integration effects of adding new resources in those same areas." [3] However, PG&E questions the appropriateness of such a proxy. PG&E believes that the cost and impact of integrating intermittent resources will increase as penetrations go up. Since the Multi-Year Analysis uses years 2002-2004 as its basis, and because renewables development can take longer than 2 years, a historical based integration cost adder will allocate only a portion of the full integration cost when it is used to evaluate future projects in the LCBF. For example, a wind project being evaluated in the 2006 RFO, and using this current Multi-Year Analysis, will incur wind integration costs associated with historical penetrations in the 2002-2004 time frame. Because this hypothetical wind project would not likely be online until 2008 or later, and would operate for 20 or more years, the historical based integration costs would not reflect the impact and cost of integrating that project in the future. As such, the CEC should incorporate forecasts into its analysis of integration cost and should expand the scope of work of the ICA to include conducting the analysis using forecasted penetration levels. Forecasts could draw on work being done in the CA Intermittency Analysis.

Staffing and completing the effort on an ongoing basis
In general, PG&E agrees with staff's recommendation to identify CPUC staff to perform the functions of an Integration Cost Analyst (ICA), which include collection/processing of necessary data components and performing periodic updates and forecast publications of integration costs for bid evaluation. Timing of these cost calculations forecasts should be commensurate with LT Planning and RPS RFO cycles. IOU's should coordinate with CAISO and ICA regarding collection, scrubbing and automation of data, which should be performed within the CAISO firewall. Workshops should be held to develop how these integration cost valuations can be practically incorporated into RPS bid evaluation process and to discuss how the ICA scope of work could be modified to accommodate methodological updates.

Next Steps
- Coordinate with CEC's CA Intermittency Analysis effort to determine cost impacts under higher levels of penetration
- Study/forecast impacts to integration costs, system reliability, operational feasibility of different levels of penetrations of intermittent resources under a range system conditions going forward. Studies should consider integration cost impacts due to:
  - ✓ Changing resource stack due to development of preferred loading order resources and retirement of aging power plants
  - ✓ Contractual constraints on ability of system to provide regulation/load following services
  - ✓ Target planning reserve margins of 15-17%
  - ✓ Over/under unit commitment due to forecasts uncertainty over weekly planning horizon

[3] P500-04-054, July 2004, California RPS Renewable Generation Integration Cost Analysis, pg 45

- ✓ Availability of ramping capability due to contractual constraints, during periods of high hydro runoff's, high maintenance periods, or future retirements of aging power plants.
- ✓ Back-up reserves
- ✓ Sensitivities that include intermittency mitigation (e.g. pump-storage, fly-wheel, forecasting, dynamic scheduling etc.).

- Results should be quantified on a monthly basis that comports with resource adequacy requirements and that could be used for market and LCBF valuations. Explore use of RA accounting in replacement of ELCC.
- MRTU: Explore "locational" impacts of RPS development
- Develop examples of how ELCC, regulation/load following metrics can be used for LCBF and overall bid evaluation.

PG&E thanks the California Energy Commission, The California Wind Energy Collaborative, The National Renewable Energy Laboratory, Oak Ridge National Laboratory, Dynamic Design Engineering, for the hard work on this report. PG&E looks forward to continuing to participate in this effort.

C.2. Response to Comments

C.2.1. CAPACITY CREDIT

PG&E recommends that the capacity credit analysis should be adapted to conform with Resource Adequacy protocols. We agree that the Integration Cost Analyst (ICA) should pursue this. One approach to bridging the gap between the ELCC and Resource Adequacy methodologies is to use the relationship between capacity factor and capacity credit, discussed in Section 2.4.2. Renewable capacity factors could be calculated monthly, based on historical delivery over the relevant periods. These results could then be priced according to the value of the relevant seasonal capacities.

This would be consistent with CaISO's statement in their revised September 25, 2005 MRTU Update (as cited in the CPUC report "Achieving 33 Percent Energy Target") that intermittent resources will likely count approximately 25-30% of their full capacity, and that the specific value will likely be tied to a historical capacity factor. While this approach does not perfectly reproduce ELCC, it has the benefit of being a simple and transparent calculation.

C.2.2. REGULATION

PG&E states that the regulation burden imposed by intermittent resources is low and "incremental costs… may get 'lost' if ISO procures AS for other system needs." The analysis technique employed fairly allocates the total regulation burden among those that cause it, regardless of size or the order that they were added to the system. The methodology correctly accounts for the system-wide reduction in regulation burden caused by aggregation, which benefits all resources, whether intermittent or not.

PG&E questions the actual total of regulation costs and volumes for CaISO. Typical control areas require between 1% and 1.5% regulation. CaISO requires a bit more because of how they use regulation resources. The analysis accounted for both the regulation that CaISO purchased and the amount that was self supplied.

PG&E suggests that there is a causal relationship between the increasing amount of regulation purchased and the decreasing cost of regulation over the study period. We are not sure that the data supports this; there were numerous other factors influencing regulation price, quantity, and energy price during the study interval. We did not try to analyze the relationship between price and volume. We do note that if PG&E is correct, then the incremental cost of regulation would be expected to decline as increasing amounts of wind and other renewables increase the amount of required regulation. We are not prepared to make that assertion at this time.

PG&E notes that relatively similar regulation costs were found for solar and for wind resources, counter to conventional wisdom. We had the same reaction while conducting the analysis. Checking the minute-to-minute time-series data, we confirmed that the variability is present in the data. The fact that the solar resources have dispatched gas co-firing may partially explain this.

PG&E notes that as the amount of wind increases the variability will increase. We agree and note that the analysis technique presented will quantify that impact.

C.2.3. LOAD FOLLOWING

PG&E points out that day-ahead scheduling protocols have changed and ESP's are now required to schedule 95% of forecasted load on a day-ahead basis. PG&E feels that this may impact the depth of the stack and increase the load following cost going forward. This is a good point. Since historic data is not yet available concerning the new scheduling rules the analysis technique we employed would not yet work. A more detailed modeling effort will be required. We suspect that there will still be significant load following capability based on the existing mix of generation technologies and their marginal costs, but that remains to be determined.

If PG&E is correct and forcing increased day-ahead scheduling does increase load following costs this has implications concerning the efficiency of the dispatch that should also be examined.

C.2.4. RAMPING CAPABILITY

PG&E correctly notes, as previously stated in the report, that the ramping analysis was limited. Proprietary information concerning actual unit capability is required to correctly assess ramping capability. As PG&E notes, information concerning contractual constraints is also required. This commercially sensitive information was not available for this public study and, should it become available in the future, should be incorporated. Note that only ramping capability of the thermal generation in California was considered in the analysis. The significant hydro ramping capability was completely ignored, providing conservative results.

C.2.5. ONGOING STUDY

As previously stated in the body of the report, ongoing study of integration costs is recommended to properly capture the effects of increasing penetration, emerging market rules, and technology advancements. Further study is also recommended as additional data – whether simulated or from new disclosures – becomes available. The adoption of an Integration Cost Analyst and the identification of a data handling entity would facilitate this recommendation.

Specifically, we agree with PG&E's statement that work in the California Intermittency Analysis Project could be very useful for future integration costs studies.

APPENDIX D: COMMENTS FROM SOUTHERN CALIFORNIA EDISON

D.1. Received Comments

The comments on the next five pages were received following the release of a draft of this report in March 2006 and a public workshop discussing the findings of this study on 3 April 2006. The analysis team would like to again thank SCE for their participation and insight throughout this study as well as for their consideration in preparing the comments herein.

SOUTHERN CALIFORNIA EDISON®

An EDISON INTERNATIONAL® Company

April 21, 2006

California Energy Commission
Attn: Cost of Integrating Renewables\[Pete Spaulding, MS-43]
1516 Ninth St.
Sacramento, CA 95814-5512
Subject: Multi-year Integration Analysis,
California Renewables Portfolio Standard Renewable Generation Integration Cost Analysis

Southern California Edison Company (SCE) hereby provides comments on the California Renewables Portfolio Standard Renewable Generation Integration Cost Analysis: Multi-Year Analysis Results and Recommendations (Analysis). As discussed below, there are many issues that have not been addressed in the Analysis which, in turn, raise concerns about the validity of the results.

I. Is California Truly Unique?

The results presented in the Analysis and in previous reports evaluating integration costs are out of line with values in the technical literature and hard data being developed as more wind projects come on line. For example, a paper presented at the American Wind Energy Association Global WindPower Conference March 28-31, 2004, in Chicago, Illinois summarized the state of the art findings at that time. Table 2 in the paper, reproduced below, shows an average value for the effect of wind power on electric power system operating costs of $3.1/MWh (3.1 mils/kWh).

TABLE 2: SUMMARY OF RESULTS

		$/MWh			
Study	Relative Wind Penetration (%)	Regulation	Load Following	Unit Commitment	Total
UWIG/Xcel	3.5	0	0.41	1.44	1.85
PacifiCorp	20	0	2.50	3.00	5.50
BPA	7	0.19	0.28	1.00 - 1.80	1.47 - 2.27
Hirst	0.06 - 0.12	0.05 - 0.30	0.70 - 2.80	na	na
We Energies I	4	1.12	0.09	0.69	1.90
We Energies II	29	1.02	0.15	1.75	2.92
Great River I	4.3				3.19
Great River II	16.6				4.53
CA RPS Phase I	4	0.17	na	na	na

The findings in the Analysis (i.e., $0.36 – 0.56/MWh) cannot be harmonized with these findings.

II. Faulty Assessment of Unit Commitment Costs

As should be obvious from the data reproduced above, the team which authored the Analysis assumed no cost ("na") for unit commitment. This assumption is erroneous, because, as discussed below, while unit commitment costs may not be visible to or observed by the ISO, the LSEs themselves provide significant unit commitment services as a result of the existence of intermittent resources on the grid, and also bear significant costs associated with such unit commitment activities.

In fact, the CAISO itself has acknowledged the additional system requirements created by intermittent resources. For example, in a presentation titled "Wind Generation Forecasting: A Balancing Authority View" made at the UWIG Fall Technical Workshop on Wind Integration held on November 7-9, 2005 in Sacramento, California, the CAISO stated:

- "California will need a portfolio of generating resources that can ramp fast, have short start up and shut down times, and have fast governor response for frequency control [to accommodate wind generation].
- "The Regulation and Load Following burden to accommodate wind generation is not trivial but can be managed with good forecasting techniques and the right mixture of generation and load resources."

These comments seem to conflict directly with the conclusions in the Analysis that there are essentially no integration costs associated with integrating wind resources. Moreover, although the CAISO does not directly acknowledge the costs associated with managing these "not trivial" phenomena, there are undeniably costs, which, if not correctly and fully incorporated into least-cost/best-fit evaluations (LC/BF) will skew bid analysis and result in a competitive disadvantage for non-intermittent renewable resources.

Not accounting for unit commitment costs is substantially inaccurate. The whole purpose of determining the integration costs for various resources is to place renewable resource alternatives on a level playing field in a competitive procurement process. Comparing an intermittent resource to a non-intermittent resource requires that the unit cost of the intermittent resource be burdened with full additional cost of "fast ramping, short start-up and shut down, and fast governor response" necessitated by the presence of the intermittent resource. Failing to make this adjustment will give an unfair competitive advantage to the intermittent resource over other, non-intermittent renewable resources. Whether the CAISO has concluded that these problems can be managed does not mean that such management occurs without cost.

III. PIRP Program Costs

SCE has also analyzed the charges assessed to SCE by the CAISO. SCE has estimated that the CAISO is charging the market 0.325 cents/kWh for the PIRP program under an allocation mechanism titled in the CAISO Tariff as the Allocation of Costs From Participating Intermittent Resources. This mechanism is intended to reallocate the un-recovered uninstructed deviation costs of the PIRP participants to the remaining scheduling coordinators participating in the CAISO. These costs have not been included in SCEs least-cost/best-fit (LCBF) analysis because in D.03-06-071 the CPUC ordered that "Intermittent resources utilize the ISO's Amendment 42 and internalize costs into bids; no further utility calculation of schedule deviations is needed…." These are real costs, however, that should be incorporated in the bid scoring process for future RPS solicitations for any bidder who intends to be in the program.

IV. Detailed Discussion of Unit Commitment Costs

The Analysis excludes unit commitment costs. ORNL/NREL have asserted, in response to informal inquiries by SCE, that the CAISO's dispatch "stack" is so deep that there should be no need to commit additional units to compensate for the intermittent nature of wind. This conclusion is incorrect.

Although *the CAISO* may not directly dispatch units or incur unit commitment costs as a result of wind resources, other stakeholders do perform this function and incur the resulting unit commitment costs. Moreover, the perceived depth of the CAISO dispatch "stack" is, in part, a result of unit commitment by stakeholders that schedule wind [not sure I follow this sentence]. Every day, SCE forecasts the expected amount of wind generation for its portfolio. In total, SCE has over 1,000 MW of nameplate wind capacity. Because of the variability of wind output, however, SCE schedules only a fraction of this capacity in any given hour. The variability is illustrated by considering that, on a portfolio basis, for the years 2002 to 2004, the standard deviation of the hourly output of SCEs wind resources is typically equal to the average of that output. SCE performs the unit commitment functions to ensure load balance after day ahead market schedules have been completed, but prior to the hour ahead market. Such units (typically thermal) are committed because of both the unpredictability of wind and the inability for the wind resources to perform in a coincident manner to reach the full aggregate nameplate rating. Often, these commitments become sunk costs, and if wind then exceeds expectations in real-time, the unit committed by SCE becomes available to the CAISO dispatch stack.

This unit commitment practice, which is necessitated entirely by the intermittent nature of wind resources, comes at a significant cost. It is reasonable to conclude that all load serving entities relying on wind are forced to have additional unit commitment. As a result, it is reasonable to use some system wide proxy for average system-wide unit commitment costs as a pricing mechanism. SCE has analyzed data available on the CAISO's OASIS site related to minimum load costs paid to must-offer units. The data shows that, for a recent 12-month period, the CAISO spent over $163 million on minimum load costs to obtain about 9.8 million MWh of capacity.[1] This results in a 12-month average cost for minimum load payment of **$16.73/MWh of capacity**. This does not include startup cost, but only minimum load costs.[2] Based on current market rules, this is a reasonable proxy for capacity-related unit commitment costs on a MWh basis. While not all of these costs are attributable to wind fluctuations, it is reasonable to include some portion of these costs as unit commitment costs. However, even if SCE were required to commit additional unit commitment capacity equivalent to 10% of the wind capacity, the unit commitment costs attributable to the wind could amount to $1.67/MWh, which is consistent with the findings in other studies.

Although these costs associated with unit commitment may not be apparent to the CAISO, they are real and must be accounted for in the bid evaluation process. The CAISO has admitted that they do not include any wind capacity in the day-ahead schedule. SCE urges the Commission to delve more thoughtfully into the facts in its Analysis.

V. Detailed Discussion of PIRP Charges

Along with other scheduling coordinators, SCE has recently been billed by the CAISO for Charge Type 721 (CT721) to clear a PIRP program balancing account. CT721 is applied to scheduling coordinators in the CAISO who are "negative deviators," or those who have over-consumed/under-delivered. SCE has been charged $235,000 since September 2004, and the entire CAISO market has been charged about $2.82 million for CT721.

[1] The data analyzed covered the period April 1, 2005 through March 17, 2006.

[2] The CAISO has indicated that the startup cost data posted on OASIS are inaccurate and so SCE did not include these data in the costs shown above.

Using estimates of production from the PIRP projects, SCE calculates that the next kWh of wind integrated into PIRP will cost the market approximately 0.325 cents/kWh. This is a "share" charge, not a direct charge. Although SCE itself does not participate in PIRP (i.e., SCE does not have a contract with an intermittent resource that participates in the program directly), it pays its share of these allocated costs, SCE's concern is that, as the intermittent program expands, these share costs, which ultimately get passed through to ratepayers as a result of the intermittent resources, will also increase. These costs have not been captured anywhere in the Analysis and yet they far eclipse the values that are being proposed to reflect the total cost of "integration."

VI. Specific Question Asked at the Workshop

SCE has provided responses to the specific questions asked by staff at the workshop in Attachment 1.

* * *

If you have any questions regarding these comments, please call me at (916) 441-2369.

Sincerely,

Manuel Alvarez

cc: Commissioner James Boyd Chairman Joseph Desmond Commissioner John L. Geesman Commissioner Jackalyne Pfannenstiel Commissioner Arthur H. Rosenfeld

Attachment 1
<u>Specific Questions Asked by CEC Staff</u>

Who should assume the responsibilities of the data handling entity? The CAISO is best situated to perform the data handling of the sensitive data that is necessary to perform these analyses in the future. This is especially true since CAISO is the owner of much of the data utilized in these analyses.

Who should assume the responsibilities of the Integration Cost Analyst? The responsibilities of Integrated Cost Analyst should be assumed by a governmental agency, and not by the IOUs.

How can the integration cost methodologies be applied to future scenarios? The integration cost methodologies need to be further refined before any assessment as to how to apply them to future scenarios can be evaluated.

How can accurate values for rated capacity be obtained? SCE interprets this question to ask what capacity value should be used for the analyses that are being performed herein. For most thermal facilities, there is the nameplate rating of the major components such as the prime mover (the turbine) and the nameplate rating of the generator. These may be the same or they may be different. The nameplate rating may even be limited by some lesser component such as the coupling between the two major components. This might be the case for geothermal, solar thermal, and biomass facilities. For wind there is the nameplate rating of the project which would be the summation of all of the wind turbines in the project. However, for all of these facilities, nameplate ratings are typically provided for specified conditions (temperature, elevation, humidity, etc.) which are not typically the case at a specific project. SCE uses the contract capacity, the capacity to which the parties agree in the contract. This has been a more satisfactory definition as a unit of determining the size of a facility over the years than the “nameplate rating” of the resource.

D.2. Response to Comments`

D.2.1. OTHER STUDIES

SCE claims that the results of this study are out of line with previous studies. This is not the case when physical and market differences in the regions studied are taken into account and when individual elements are examined. Major recent studies, for example, have all found similar regulation impacts as analysis techniques have matured. Market structure has an important impact on some studies. The Xcel Minnesota study, for example, found significant costs that were tied to the lack of an hourly energy market in the Midwest coupled with a specific generation mix with few intermediate cost units.

D.2.2. UNIT COMMITMENT

SCE comments that impacts on unit commitment costs were not considered in the analysis. The analysis addressed unit commitment within the data and scope constraints by examining the depth of the available generation stack. Over the historic period studied there was significant stack depth implying significant dispatch flexibility. The impact of renewable variability on the stack depth was small by comparison.

While SCE discusses the impact of wind forecast errors on unit commitment, SCE does not address the impact of load forecast errors or the interaction of load and wind forecast errors. Load forecast errors are significantly greater than wind forecast errors if only because there is so much more load. Fortunately, the two forecast errors benefit from aggregation. The fact that load serving entities found it advantageous to voluntarily significantly underschedule and use the stack provides some indication that the generation mix in California provides significant flexibility.

We certainly agree that additional analysis would be useful.

D.2.3. PIRP

PIRP is the Participating Intermittent Resource Program offered through the CA ISO. This protocol allows intermittent resources, such as wind-powered generators and other resources with an uncontrollable fuel source, to schedule energy in the ISO forward market without incurring imbalance charges when the delivered energy differs from the scheduled amount. For more detailed information, go to:
http://www.caiso.com/docs/2003/01/29/2003012914230517586.html

The cost specifics of the CaISO PIRP are beyond the scope of this study; however, we note that PIRP costs are directly tied back to CaISO's forecasting program and include various charges for administration. SCE's questions regarding their PIRP costs have been forwarded to CaISO.

Should participation in PIRP significantly increase, we recommend that the ICA investigate the specifics of it further.

D.2.4. ONGOING STUDY

As previously stated in the body of the report, ongoing study of integration costs is recommended to properly capture the effects of increasing penetration, emerging market rules, and technology advancements. Further study is also recommended as additional data – whether simulated or from new disclosures – becomes available. The adoption of an Integration Cost Analyst and the identification of a data handling entity would facilitate this recommendation.